

Lecture Notes in Computer Science

1509

Colin P. Williams (Ed.)

Quantum Computing and Quantum Communications

First NASA International Conference, QCQC'98
Palm Springs, California, USA, February 1998
Selected Papers



Springer

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo

Colin P. Williams (Ed.)

Quantum Computing and Quantum Communications

First NASA International Conference, QCQC'98
Palm Springs, California, USA
February 17-20, 1998
Selected Papers



Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany

Juris Hartmanis, Cornell University, NY, USA

Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editor

Colin P. Williams

Jet Propulsion Laboratory, California Institute of Technology

Quantum Algorithms and Technologies Group

Information and Computing Technologies Research Section

Mail Stop 126-347, Pasadena, CA 91109-8099, USA

E-mail: Colin.P.Williams@jpl.nasa.gov

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Quantum computing and quantum communications : first NASA international conference ; selected papers / QCQC'98, Palm Springs, California, USA, February 17 - 20, 1998. Colin P. Williams (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 1999

(Lecture notes in computer science ; Vol. 1509)

ISBN 3-540-65514-X

CR Subject Classification (1998): F.1, E.3, E.4, F.2

ISSN 0302-9743

ISBN 3-540-65514-X Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1999

Printed in Germany

Typesetting: Camera-ready by author

SPIN 10692655 06/3142 - 5 4 3 2 1 0

Printed on acid-free paper

Preface

Colin P. Williams*

Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, CA 91109-8099,
email: Colin.P.Williams@jpl.nasa.gov

Over the past half century computers have gone from being the room-sized servants of a privileged few to the totable companions of business travellers, school children, and just about anyone who can point and click a mouse. In part, this transformation was made possible by the dramatic miniaturization in the basic components of a computer. This trend was quantified in 1964 by Gordon Moore, one of the founders of Intel, who noticed that the amount of information that could be stored on a given amount of silicon doubled roughly every 18 months. The doubling trend continues to this day and, by crude extrapolation, predicts that the computers of 2020 might be approaching the one-atom-per-bit level.

Physical systems such as atoms, however, behave in ways that are very different from everyday objects. In fact they are governed by the laws of quantum mechanics rather than classical mechanics. In the early 1980's some foresighted physicists, such as Charles Bennett (our conference Chairperson), Rolf Landauer, Paul Benioff, David Deutsch, and Richard Feynman, began to question what it would mean for a computer to operate at the one-atom-per-bit scale. The elementary operations of such a computer would need to be described in terms of quantum mechanics. Recently, physicists and computer scientists have come to appreciate that certain quantum effects, in particular superposition, interference, entanglement, non-locality, indeterminism, and non-clonability, allow entirely new kinds of tasks to be performed. These tasks include teleporting quantum information, establishing shared secret cryptographic keys, searching unstructured "virtual" databases, factoring composite integers, simulating physical systems, and enhancing the capacity of classical communication channels.

In an effort to elucidate new ideas, and to push the envelope on existing ones, in 1998 my colleagues, Leon Alkalai, Richard Doyle, Amir Fijany, Sandeep Gulati, Benny Toomarian, Michail Zak, and I at the Jet Propulsion Laboratory, organized the "First NASA International Conference on Quantum Computing and Quantum Communications", in Palm Springs, California. We imposed an a priori structure of the meeting that was designed to solicit contributions on how quantum computing might impact NASA mission objectives in computation and communications. I am delighted to say that NASA obtained great value for money from this conference. The papers contained in this volume are a testament to the rich diversity of ideas that were presented.

* Supported by the NASA/JPL Center for Integrated Space Microsystems, NASA Advanced Concepts Office and the NASA Information and Computing Research Technologies Program

NASA's interest in quantum computing and quantum communications stems from its need to solve daunting computational and communications problems. In particular, spacecraft design, mission planning, observation scheduling, design optimization, image processing, data assimilation, robotic vision, all impose extreme demands on computational resources. Given that quantum computers are known to speed up the solution of some computational problems and facilitate more efficient communications, we'd like to circumscribe their capabilities on the computational and communications problems of interest to NASA. In this regard, the Palm Springs conference was a tremendous success with several new ideas emerging for tackling structured search problems, Earth-to-space quantum key distribution, improved precision atomic clocks, and quantum gyroscopy.

The papers appearing in this volume are organized by the following five themes: *Entanglement and Quantum Algorithms*, *Quantum Cryptography*, *Quantum Copying and Quantum Information Theory*, *Quantum Error Correction and Fault-Tolerant Quantum Computing*, and *Embodiments of Quantum Computers*. With such a diverse range of contributions we hope that there will be something in this volume to interest everyone.

I would like to thank the program committee and reviewers who all contributed greatly to the success of the NASA QCQC'98 conference. I pay special thanks to Dr. Charles Bennett of IBM for serving as the conference Chairperson. Finally, I would like to thank our NASA/JPL sponsors who supported this conference both financially and intellectually. In particular, I thank the NASA/JPL Center for Integrated Space Microsystems (CISM), the Ballistic Missile Defense Organization, the NASA/JPL Center for Space Microelectronics Technology (CSMT), and the NASA Autonomy and Information Technology Program Office.

Colin P. Williams

Table of Contents

Entanglement and Quantum Algorithms

Multi-particle Entanglement via Two-Particle Entanglement	1
<i>G. Brassard and T. Mor</i>	
Quantum Wavelet Transforms: Fast Algorithms and Complete Circuits . . .	10
<i>A. Fijany and C.P. Williams</i>	
Quantum Computer for Fluid Simulation	34
<i>J. Yepez</i>	
Quantum Entanglement and the Communication Complexity of the Inner Product Function	61
<i>R. Cleve, W. van Dam, M. Nielsen, and A. Tapp</i>	
Quantum Recurrent Networks for Simulating Stochastic Processes	75
<i>M. Zak and C.P. Williams</i>	
Correlation Between Correlations: Process and Time in Quantum Networks	89
<i>G. Mahler and I. Kim</i>	
Quantum Effects in Algorithms	103
<i>R. Jozsa</i>	
Automated Design of Quantum Circuits	113
<i>C.P. Williams and A.G. Gray</i>	
Quantum Search on Structured Problems	126
<i>L.K. Grover</i>	
Generalized Grover Search Algorithm for Arbitrary Initial Amplitude Distribution	140
<i>D. Biron, O. Biham, E. Biham, M. Grassl, and D. Lidar</i>	
Quantum Database Search by a Single Query	148
<i>D.P. Chi and J. Kim</i>	
Quantum Computer Cannot Speed Up Iterated Applications of a Black Box	152
<i>Y. Ozhigov</i>	
Quantum Resonance for Solving NP-complete Problems by Simulations . . .	160
<i>M. Zak</i>	

Computational Complexity and Physical Law	167
<i>D. Abrams and S. Lloyd</i>	

The Hidden Subgroup Problem and Eigenvalue Estimation on a Quantum Computer	174
<i>M. Mosca and A. Ekert</i>	

A Diakoptic Approach to Quantum Computation	174
<i>G. Castagnoli and D. Monti</i>	

Quantum Cryptography

Practical Free-Space Quantum Cryptography	200
<i>R.J. Hughes, W.T. Buttler, P.G. Kwiat, S.K. Lamoreaux, G.G. Luther, G.L. Morgan, J.E. Nordholt, C.G. Peterson, and C.M. Simmons</i>	

Quantum Cryptography, Eavesdropping and Unsharp Spin Measurement . .	214
<i>S. Roy and G. Kar</i>	

Quantum Copying and Quantum Information Theory

Information-Theoretic Aspects of Quantum Copying	218
<i>N.J. Cerf</i>	

Universal Optimal Cloning of Qubits and Quantum Registers	235
<i>V. Bužek and M. Hillery</i>	

Entanglement of Assistance	247
<i>D.P. DiVincenzo, C.A. Fuchs, H. Mabuchi, J.A. Smolin, A. Thapliyal, and A. Uhlmann</i>	

What Information Theory Can Tell Us about Quantum Reality	258
<i>C. Adami and N.J. Cerf</i>	

Quantum Generalization of Conditional Entropy and Information	269
<i>L.B. Levitin</i>	

Accessible Information in Multi-access Quantum Channels	276
<i>A.E. Allahverdyan and D.B. Saakian</i>	

Capacities of Quantum Channels and Quantum Coherent Information . . .	285
<i>M. Westmoreland and B. Schumacher</i>	

Strengthened Lindblad Inequality: Applications in Non-equilibrium Thermodynamics and Quantum Information Theory	296
<i>D.B. Saakian and A.E. Allahverdyan</i>	

Quantum Error Correction and Fault-Tolerant Quantum Computing

Fault-Tolerant Quantum Computation with Higher-Dimensional Systems . .	302
<i>D. Gottesman</i>	
Quantum Convolution Error Correction Codes	314
<i>H.F. Chau</i>	
On the Existence of Nonadditive Quantum Codes	325
<i>V.P. Roychowdhury and F. Vatan</i>	
Quantum Error Correction Is Applicable for Reducing Spatially Correlated Decoherence	337
<i>L.-M. Duang and G.-C. Guo</i>	
Topological Quantum Computation	341
<i>R.W. Ogburn and J. Preskill</i>	

Embodiments of Quantum Computers

NMR GHZ	357
<i>R. Laflamme, E. Knill, W.H. Zurek, P. Catasti, S.V.S. Mariappan</i>	
Quantum Computing Using Electron-Nuclear Double Resonances	364
<i>C.M. Bowden, J.P. Dowling, and S.P. Hotaling</i>	
Physical Implementations for Quantum Communication in Quantum Networks	373
<i>H.-J. Briegel, J.I. Cirac, W. Dur, S.J. van Enk, H.J. Kimble, H. Mabuchi, and P. Zoller</i>	
An Optical Approach to Quantum Computing	383
<i>J.D. Franson and T.B. Pittman</i>	
Quantum Computation with Linear Optics	391
<i>C. Adami and N.J. Cerf</i>	
Decoherence Control for Optical Qubits	402
<i>D. Vitali and P. Tombesi</i>	
Adiabatic Controlled-NOT Gate for Quantum Computation	413
<i>D.V. Averin</i>	
Trapped Ion Quantum Computer Research at Los Alamos	426
<i>D.F.V. James, M.S. Gulley, M.H. Holzschneider, R.J. Hughes, P.G. Kwiat, S.K. Lamoreaux, C.G. Peterson, V.D. Sandberg, M.M. Schauer, C.M. Simmons, D. Tupa, P.Z. Wang, and A.G. White</i>	

Arrays of Elliptical Ion Traps for Parallel Quantum Computing 438
R.G. DeVoe

Simulating the Effect of Decoherence and Inaccuracies on a Quantum
Computer 447
K.M. Obenland and A.M. Despain

Implementation of Quantum Controlled-NOT Gates Using Asymmetric
Semiconductor Quantum Dots 460
G. Brassard

Spatiotemporal Dynamics of Quantum Computing Solid Dipole-Dipole
Block Systems 468
H. Matsueda

Author Index 479

Multi-particle Entanglement via Two-Particle Entanglement

Gilles Brassard* and Tal Mor**

Université de Montréal, Département IRO
C.P. 6128, succursale centre-ville, Montréal (Québec), Canada H3C 3J7
{brassard,mor}@iro.umontreal.ca

Abstract. Entanglement between n particles is a generalization of the entanglement between two particles, and a state is considered entangled if it cannot be written as a mixture of tensor products of the n particles' states. We present the key notion of *semi-separability*, used to investigate n -particle entanglement by looking at two-subsystem entanglement between its various subsystems. We provide *necessary conditions* for n -particle separability (that is, *sufficient conditions* for n -particle entanglement). We also provide necessary and sufficient conditions in the case of pure states. By a surprising example, we show that such conditions are not sufficient for separability in the case of mixed states, suggesting entanglement of a strange type.

1 Introduction

Entanglement between two particles A and B provides correlations that have no classical counterpart [5, 2, 9]. A two-particle pure state can either be a tensor product of the one-particle states, or else it contains entanglement between them. For mixed states, there is a third possibility, which is a mixture of tensor-product states. Such a state is not entangled, and the name *separable state* is used to consider both tensor products and mixtures of tensor products. Fulfilling Bell's inequalities provides necessary conditions for separability. (Equivalently, breaking Bell's inequalities provides sufficient conditions for entanglement.) There are many others necessary conditions for separability, such as the inability for the state to be used to teleport a qubit [3]. The other direction is much more subtle; necessary conditions for 2-particle entanglement (that is, sufficient conditions for separability) were given in some special cases [10], but are still missing in the general two-particle case [7].

For three and more particles, the situation is more complex as we demonstrate in this paper. Entanglement between n particles is a generalization of the entanglement between two particles, and a state is considered entangled if it cannot be written as a mixture of tensor products of the n particles' states. We discuss properties which, by definition, are non-trivial only for systems composed of

* Supported in part by Canada's NSERC, Québec's FCAR, and the Canada Council.

** Supported in part by Canada's NSERC and Québec's FCAR.

more than two particles: two-subsystem entanglement or separability between the various possible subsystems. (There are six possibilities in the three-particle case.)

We define *k*-subsystem semi-separability of a system as follows: Let a system contain n particles such that $n > k$. Divide all particles into k or $k + 1$ non-empty groups of particles (k or $k + 1$ subsystems). If we divided into $k + 1$ subsystems, trace out one of them to be left with k subsystems in any case. The original system is semi-separable if those k subsystems are separable.

We shall be interested in two more specific cases, depending on whether we trace out one subsystem or not.

***k*-subsystem total semi-separability** Divide all particles into k non-empty groups (k subsystems) and check separability of the k subsystems.

***k*-subsystem partial semi-separability** Divide all particles into $k + 1$ non-empty groups ($k + 1$ subsystems), trace out one subsystem, and check separability of the remaining k subsystems.

We shall refer to these as *k*-TSS and *k*-PSS, respectively. The case $k = 2$ is of special interest since separability of two particles (two subsystems in our case) has been extensively studied. Note that these notions (when $k > 2$) can be reduced to two subsystems semi-separability. We concentrate on $k = 2$ in the following and drop the index k unless there is a danger of confusion.

These properties provide a new insight into many-particle entanglement and allow us to find a surprising type of entanglement, which shows a new fundamental difference between the properties of pure states and of mixed states.

We use these properties to provide a partial classification of n -particle entanglement and separability in terms of the much simpler (albeit still only partially solved) problem of separability of two subsystems, each possibly composed of several of the original particles. We provide *necessary conditions* for n -particle separability. We also provide necessary and sufficient conditions in the case of pure states. These properties may also be useful in the future in providing a complete classification of separability versus entanglement of many particles. Many-particle entanglement was recently discussed in other works [11, 8, 4].

2 Conditions for Multi-particle Separability

Most of the discussion in this section is restricted (for simplicity) to two-subsystem entanglements in a system composed of three particles, but is true for two subsystems in a system composed of n particles unless explicitly stated otherwise. [Possibly, the discussion is also true for $k > 2$.] Furthermore, we restrict our examples to qubits, to make it simple and clear, but *all* the *Facts* we provide apply to higher-dimensional Hilbert spaces than qubits.

Let Alice, Bob and Carol (who are spatially separated) have three qubits (denoted by A , B and C respectively), prepared in some joint pure state $|\Psi^{(r)}\rangle$, where (r) is used to index one of possibly many such states:

$$|\Psi^{(r)}\rangle = \sum_{i=0}^7 \alpha_i^{(r)} |i\rangle = \alpha_{000}^{(r)} |0_A 0_B 0_C\rangle + \alpha_{001}^{(r)} |0_A 0_B 1_C\rangle + \cdots + \alpha_{111}^{(r)} |1_A 1_B 1_C\rangle, \quad (1)$$

with $\sum_i |\alpha_i|^2 = 1$. The states $|0\rangle$ and $|1\rangle$ are basis vectors of each qubit, and $|00\rangle \equiv |0\rangle \otimes |0\rangle$. We usually avoid using the tensor product notation \otimes unless we want to emphasize it. The most general three-qubit state is a mixture of states of this type

$$\rho = \sum_r p_r |\Psi^{(r)}\rangle \langle \Psi^{(r)}|, \quad (2)$$

with $\sum_r p_r = 1$. Any such state can be written in various forms. A three-particle state is considered separable, or non-entangled (NE), if and only if it can be written as a mixture of tensor products

$$\rho_{NE} = \sum_s p_s [\rho_A^{(s)} \otimes \rho_B^{(s)} \otimes \rho_C^{(s)}]. \quad (3)$$

A separable *pure state* can be presented using its Schmidt decomposition [9], provided we separate one subsystem at a time (and do it recursively if there are more particles). Then it is necessarily in a tensor product of one-particle states $|\Psi_{NE}\rangle = |\Psi_{AB}\rangle \otimes |\phi_C\rangle = |\phi_A\rangle \otimes |\phi_B\rangle \otimes |\phi_C\rangle$ where the $|\phi\rangle$'s are one qubit states.

Let us study a three-particle entangled state by looking at the entanglement between its various two subsystems. There are three options for partial semi-separability: tracing out (ignoring) one particle to be left with two particles. Similarly, there are three options for total semi-separability: considering two particles as one subsystem, to be left with this combined subsystem and the remaining particle. If we trace out particle C and the remaining state of systems A and B is separable, we denote this 2-subsystem partial semi-separability by $PSS_C(A; B)$. If the subsystem composed of AB is separable from C , we denote this two-subsystem total semi-separability by $TSS(AB; C)$. Similar notation can be written for the other four options obtained by cyclic permutations of the three particles. A negation is denoted by $\overline{TSS}(AB; C)$ saying that AB is entangled with C . In the general case of n particles and k subsystems, more terms could appear, for instance 4- $TSS(A; BC; D; E)$ and 3- $PSS_{CD}(A; B; E)$. Once these new notions are established, we can use them to prove some simple facts, relating n -particle entanglement to the better investigated entanglement between two systems.

A separable state [such as in Equation (3)] presents all three possible TSS properties. This is immediately obtained from the fact that the state is still separable when collecting two particles together: for instance, collecting A and B together yields $\rho_{NE} = \sum_s p_s [\rho_{AB}^{(s)} \otimes \rho_C^{(s)}]$. Thus,

Fact 1 If a state does not present *all* cases of TSS , then it is entangled.

In other words, the existence of all possible TSS is a necessary condition for separability (and the existence of one \overline{TSS} is a sufficient condition for entanglement). If one pair of subsystems is entangled (even if other pairs are not) the

entire system is entangled. If none of the possible pairs is entangled the system *might* be separable. If each pair of subsystems satisfies a necessary condition for separability (for instance, it cannot be purified to a state which violates Bell's inequalities) but is not *known* to be separable, then the entire system satisfies a necessary condition for separability.

We conjecture that the existence of all possible *TSS* is also a sufficient condition for separability, but this important question is still left open in the general case of mixed states.

A separable state [such as in Equation (3)] satisfies

$$\text{Tr}_C \rho_{NE} = \sum_s p_s \text{Tr}_C [\rho_A^{(s)} \otimes \rho_B^{(s)} \otimes \rho_C^{(s)}] = \sum_s p_s [\rho_A^{(s)} \otimes \rho_B^{(s)}] \quad (4)$$

due to the linearity of the trace-out operation. Similar equations can be written if *A* or *B* are traced out. Therefore it presents all *PSS* properties. Thus,

Fact 2 If a state does not present all cases of *PSS*, then it is entangled.

In other words, the existence of all possible *PSS* is a necessary condition for separability (and the existence of one *PSS* is a sufficient condition for entanglement). The converse is not true and a three-particle *entangled state* might present all *PSS* properties: the GHZ–Mermin state [6] is a good example—see Section 3. Therefore, the existence of *all PSS* is a necessary condition for separability, but not a sufficient one.

Note that a similar argument implies that if a state fulfills *TSS*(*A*; *BC*) then automatically it also fulfills *PSS*_{*C*}(*A*; *B*) and *PSS*_{*B*}(*A*; *C*), and similar conclusions can be obtained by cyclic permutations.

For pure states, we now obtain conditions that are both necessary and sufficient for separability. (We need only prove that they are sufficient, since the fact that they are necessary is already shown using the previous facts.) We omit Dirac's bracket notation unless confusion could arise.

Fact 3 A pure state that presents all *TSS* properties (for all possible decompositions of two subsystems) is separable.

Proof. We prove it in the case of three particles and leave the general case for the final paper. If the three-particle pure state is totally semi-separable for all possible ways of decomposing the subsystems then it can be written as $\Psi_{ABC} = \Psi_{AB} \otimes \Psi_C$. If Ψ_{AB} can be written as $\Psi_A \otimes \Psi_B$ then $\Psi_{ABC} = \Psi_A \otimes \Psi_B \otimes \Psi_C$ and the state is separable.

Since Ψ_{ABC} presents all *TSS* properties it can also be written as $\Psi_{ABC} = \Phi_A \otimes \Phi_{BC}$. Assuming for a contradiction that Ψ_{AB} cannot be written as $\Psi_A \otimes \Psi_B$, it can be decomposed (due to Schmidt decomposition [9]) as $\alpha \Psi_A \otimes \Psi_B + \beta \Psi'_A \otimes \Psi'_B$ with $\alpha \neq 0$ and $\beta \neq 0$, where χ' denotes an orthogonal state to χ for any state χ . Now $\Psi_{ABC} = \alpha \Psi_A \otimes (\Psi_B \otimes \Psi_C) + \beta \Psi'_A \otimes (\Psi'_B \otimes \Psi_C)$ is (by treating *B* and *C* together) the Schmidt decomposition of *A* and *BC*, showing entanglement between *A* and *BC*, in contradiction to $\Psi_{ABC} = \Phi_A \otimes \Phi_{BC}$. \square

Fact 3 can be strengthened. In the case of three particles, we have:

Fact 3' A three-particle pure state that is *TSS* *twice* is separable.

Proof. Without loss of generality, suppose that the two *TSS* properties are $TSS(AB; C)$ and $TSS(A; BC)$. We can make use of the proof of Fact 3 since it used only these two semi-separabilities anyhow. \square

Fact 4 A pure state that presents a particular *TSS* in which one particle is separable from all the others, and also presents all *PSS*, is separable.

Proof. Again, we prove it in the case of three particles and leave the general case for the final paper. Since the three-particle pure state is totally semi-separable once [say, $TSS(AB; C)$] it can be written as $\Psi_{ABC} = \Psi_{AB} \otimes \Psi_C$. We now trace out particle C to get $\Psi_{AB} = \Psi_A \otimes \Psi_B$ since the state presents all *PSS*. Thus, $\Psi_{ABC} = \Psi_A \otimes \Psi_B \otimes \Psi_C$. \square

We conclude that Facts 3, 3' and 4 provide necessary and sufficient conditions for the separability of pure states.

Surprisingly, Fact 4 is not true for mixed states. We present in Section 4 a three-particle *mixed state* that presents one case of *TSS* [say, $TSS(A; BC)$] and all *PSS*, yet it is entangled.

3 A Few Simple Examples

Let us first provide some examples of pure states in order to obtain a better intuition about the meaning of the different semi-separability conditions. We also describe the techniques required for verifying the existence of semi-separabilities. Although such tools are not used here, they might be required for a more complete analysis of the possible connections between semi-separability of mixed states.

Our first example is

$$|\Psi_{En}\rangle = |\Psi_A\rangle \otimes |\Psi_{BC}^-\rangle \quad (5)$$

with $|\Psi^-\rangle = (1/\sqrt{2})[|01\rangle - |10\rangle]$ is the singlet state of two qubits. When the non-entangled particle is traced out, the other two are still entangled so we have $\overline{PSS}_A(B; C)$, but when B or C are traced out the remaining particles are not entangled, so we have $PSS_B(A; C)$ and $PSS_C(A; B)$. Clearly it is also $TSS(A; BC)$, $\overline{TSS}(AB; C)$ and $\overline{TSS}(AC; B)$: When the two particles A and B are considered as one subsystem AB , it is fully entangled with subsystem C since the state can be written as $|\Psi_A 0\rangle \otimes |1\rangle - |\Psi_A 1\rangle \otimes |0\rangle$, and the states $|\Psi_A 0\rangle$ and $|\Psi_A 1\rangle$ play the role of the relevant two basis vectors $|0\rangle$ and $|1\rangle$ of the four-dimensional subsystem.

Consider a three-qubit case. If we write the three-particle general pure state as

$$|\Psi\rangle = |0_A\rangle \left(\sum_{k=0}^3 \alpha_{0k} |k_{BC}\rangle \right) + |1_A\rangle \left(\sum_{k=0}^3 \alpha_{1k} |k_{BC}\rangle \right), \quad (6)$$

tracing out subsystem A shall leave the other two subsystems in a (possibly mixed) state

$$\begin{aligned} \rho &= \left(\sum_{k=0}^3 \alpha_{0k} |k_{BC}\rangle \right) \left(\sum_{l=0}^3 \alpha_{0l}^* \langle l_{BC}| \right) + \left(\sum_{k=0}^3 \alpha_{1k} |k_{BC}\rangle \right) \left(\sum_{l=0}^3 \alpha_{1l}^* \langle l_{BC}| \right) \\ &= \sum_{k=0}^3 \sum_{l=0}^3 (\alpha_{0k} \alpha_{0l}^* + \alpha_{1k} \alpha_{1l}^*) |k\rangle \langle l|. \end{aligned} \quad (7)$$

Using the transposition technique recently discovered by Peres [10], it is easy to check whether or not this reduced density matrix is separable for any particular case. This result (and similar results for tracing out the other particles) can be used to verify if entanglement between two particles exists. For larger systems (such as qutrits, which are 3-dimensional Hilbert spaces), no perfect way to tell if the two remaining qutrits are separable have been found so far, and it is known that Peres's criterion does not apply [7].

To verify that all cases of PSS are present, the three possibilities must be checked. For instance, the state

$$|\Psi_{ABC}\rangle = \frac{\cos \theta}{\sqrt{2}} |000\rangle + \frac{\cos \theta}{\sqrt{2}} |011\rangle + \frac{\sin \theta}{\sqrt{2}} |100\rangle - \frac{\sin \theta}{\sqrt{2}} |111\rangle \quad (8)$$

(with $|000\rangle \equiv |0_A 0_B 0_C\rangle$) leads to the reduced density matrix

$$\rho_{BC} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & \cos 2\theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \cos 2\theta & 0 & 0 & 1 \end{pmatrix} \quad (9)$$

which is entangled for $\theta \neq \pi/4$.

Checking the TSS property is sometimes impossible even for 3 qubits, using the best techniques currently available. If we write the three-particle general pure state of Equation (6) using a Schmidt decomposition, it can be written as

$$|\Psi\rangle = \cos \theta |\phi_A\rangle |\chi_{BC}\rangle + \sin \theta e^{i\phi} |\phi'_A\rangle |\chi'_{BC}\rangle \quad (10)$$

(where $|\psi'\rangle$ stands for a state orthogonal to $|\psi\rangle$). Although it is not obvious to find this decomposition, this can always be done [9]. When both terms appear in the Schmidt decomposition, the state is $\overline{TSS}(A; BC)$. Generalizing this argument to mixed states is usually impossible with the current techniques. Peres's transposition [10] works unconditionally [7] only for two qubits (2×2) and for a qubit and a qutrit (2×3), and it can only be used to check entanglement of a four-dimensional system with a two-dimensional system, but not to check separability.

For the mixed state that is of interest to us, it is possible to find the appropriate Schmidt decomposition as we shall see in Section 4.

In the following examples of pure states, the state is already given in such a Schmidt decomposition (for all possible combinations of pairs), so they clearly present no *TSS*. For a GHZ–Mermin state [6]

$$\Psi_{GHZM} = (1/\sqrt{2})[|0_A 0_B 0_C\rangle + |1_A 1_B 1_C\rangle], \quad (11)$$

let us consider the particles of Bob and Carol as one four-dimensional subsystem. Then this state becomes a Bell state

$$(1/\sqrt{2})[|0_A(00)_{BC}\rangle + |1_A(11)_{BC}\rangle] \quad (12)$$

so that $|00\rangle$ and $|11\rangle$ play the role of $|0\rangle$ and $|1\rangle$, the relevant two basis vectors of the four-dimensional subsystem. Clearly, the same is true for the other two cases. This state presents all *PSS*: when one particle is traced out, the other two are in a mixture-of-product state. Similar arguments apply to the following state:

$$\Psi_{Zei} = (1/\sqrt{2})[|0_A \Psi_{BC}^+\rangle + |1_A \Psi_{BC}^-\rangle], \quad (13)$$

with $|\Psi^+\rangle = (1/\sqrt{2})[|01\rangle + |10\rangle]$. Particle A is entangled to the subsystem composed of BC together. By transforming to the $|0 \pm 1\rangle$ basis of particle A , this state becomes a GHZ–Mermin state, so it presents all properties as above.

As we previously said, three-particle entangled *pure* states can at most present one *TSS*. We conjecture that this is also true for a mixed entangled state, but we leave this question open¹ for future research.

4 A Surprising Example

We now present our surprising example: a three-particle mixed state that presents one *TSS* and all *PSS*, but yet is entangled! [This should be contrasted with Fact 4.]

Consider the state that is composed of an equal mixture of the two states $|\Psi_1\rangle = |0_A\rangle \otimes |\Psi_{BC}^+\rangle$ and $|\Psi_2\rangle = |1_A\rangle \otimes |\Psi_{BC}^-\rangle$, and thus

$$\rho = \frac{1}{2} [(|0_A\rangle\langle 0_A|) \otimes (|\Psi_{BC}^+\rangle\langle \Psi_{BC}^+|) + (|1_A\rangle\langle 1_A|) \otimes (|\Psi_{BC}^-\rangle\langle \Psi_{BC}^-|)]. \quad (14)$$

This state presents all *PSS*: when particle A is traced out, the other particles are left in an equal mixture of the two Bell states $|\Psi^+\rangle$ and $|\Psi^-\rangle$, which is separable since this is the same as an equal mixture of $|01\rangle$ and $|10\rangle$; when particle B (resp. C) is traced out, it is clear that the particle entangled to it, C (resp. B), does not become entangled with A . This state also presents *TSS*($A; BC$): when subsystems B and C are considered as one subsystem, A is clearly separable from it since they are written as mixture of products.

To prove that it is an entangled state, let us show that it is not semi-separable for *TSS*($AB; C$). Then, Fact 1 implies that the three-particle state

¹ The analysis is much more complicated since semi-separability for mixed state does not mean tensor product but mixture-of-tensor-products.

is entangled. The two pure states that are mixed to give ρ can be written as $|\Psi_1\rangle = |0_A 0_B\rangle \otimes |1_C\rangle + |0_A 1_B\rangle \otimes |0_C\rangle$ and $|\Psi_2\rangle = |1_A 0_B\rangle \otimes |1_C\rangle - |1_A 1_B\rangle \otimes |0_C\rangle$, up to normalization. Now we can see that C is entangled to different two-dimensional subspaces of AB in each of these pure states, thus mixing cannot reduce or cancel this entanglement.

Such a state is surprising: although Alice is not entangled with the subsystem of Bob and Carol (together) she can *control* their entanglement. If she measures in the $|0 \pm 1\rangle$ basis, Bob and Carol will not be entangled whatsoever, but if she measures in the computational basis, Bob and Carol are entangled without knowing it, and their state depends on Alice's measurement result. Thus, Bob and Carol are in a separable state if they ignore (trace out) Alice's knowledge, but they become entangled once they receive Alice's result using classical communication.

5 Generalizations and Conclusions

To summarize, we analysed three-particle entanglement/separability (and beyond) in terms of its possible two-subsystems entanglements/separability, which we call total and partial semi-separability. We presented necessary conditions for separability and also sufficient conditions in the case of pure states. We also discussed possible generalizations and presented a surprising state.

Generalizations to larger systems are more complicated. A four-qubit system A, B, C and D , can be discussed in terms of two-particle/three-particle entanglement, and semi-separability to $2\text{-}TSS(A;BCD)$ etc., and $2\text{-}TSS(AB;CD)$ etc., and $3\text{-}TSS(A;B;CD)$ etc., and also $3\text{-}PSS_A(B;C;D)$ etc., and $2\text{-}PSS_{AB}(C;D)$ etc., where all the "etc." refer to permutations of the particles. A relatively simple (and yet interesting) example can be built by replacing $|0_A\rangle$ and $|1_A\rangle$ of the previous example by two Bell states of two particles (A and D):

$$\rho = \frac{1}{2} [(|\Psi_{AD}^+\rangle\langle\Psi_{AD}^+|) \otimes (|\Psi_{BC}^+\rangle\langle\Psi_{BC}^+|) + (|\Psi_{AD}^-\rangle\langle\Psi_{AD}^-|) \otimes (|\Psi_{BC}^-\rangle\langle\Psi_{BC}^-|)] \quad (15)$$

so that Bob and Carol are entangled if Alice and David (D) measure their entanglement, but if Alice and David measure in a product basis, Bob and Carol become disentangled.

Acknowledgements

We are grateful to Dorit Aharonov, Charles H. Bennett, Isaac Chuang, David DiVincenzo, Pawel Horodecki, Asher Peres, Peter Shor and Bill Wootters for stimulating discussions on these topics. Part of this work was completed during the 1997 Elsag-Bailey – I.S.I. Foundation research meeting on quantum computation.

References

1. D. Aharonov and M. Ben-Or, “Polynomial simulations of decohered quantum computers”, *Proceedings of 37th Annual IEEE Symposium on the Foundations of Computer Science*, 46–55 (1996).
2. J. S. Bell, “On the Einstein Podolsky Rosen paradox”, *Physics* **1**, 195–200 (1964).
3. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels”, *Physical Review Letters* **70**, 1895–1899 (1993).
4. R. Clifton, D. V. Feldman, M. L. G. Redhead and A. Wilce, “Hyperentangled States”, Los Alamos archive (<http://xxx.lanl.gov>): quant-ph 9711020.
5. A. Einstein, B. Podolsky and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?”, *Physical Review* **47**, 777–780 (1935). Reprinted in *Quantum Theory and Measurement*, J. A. Wheeler and W. Z. Zurek, Editors, Princeton University Press, 1983.
6. D. M. Greenberger, M. A. Horne and A. Zeilinger, “Multiparticle interferometry and the superposition principle”, *Physics Today* **46**, 22–29 (1993).
7. M. Horodecki, P. Horodecki and R. Horodecki, “Separability of mixed states: Necessary and sufficient conditions”, Los Alamos archive: quant-ph 9605038.
8. N. Linden and S. Popescu, “On multi-particle entanglement”, Los Alamos archive: quant-ph 9711016.
9. A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers, Dordrecht (1993).
10. A. Peres, “Separability criterion for density matrices”, *Physical Review Letters* **77**, 1413–1415 (1996).

Quantum Wavelet Transforms: Fast Algorithms and Complete Circuits

Amir Fijany and Colin P. Williams

Jet Propulsion Laboratory, California Institute of Technology
MS 126-347, 4800 Oak Grove Drive
Pasadena, CA 91109

Amir.Fijany@jpl.nasa.gov and Colin.P.Williams@jpl.nasa.gov

Abstract. The quantum Fourier transform (QFT), a quantum analog of the classical Fourier transform, has been shown to be a powerful tool in developing quantum algorithms. However, in classical computing there is another class of unitary transforms, the wavelet transforms, which are every bit as useful as the Fourier transform. Wavelet transforms are used to expose the multi-scale structure of a signal and are likely to be useful for quantum image processing and quantum data compression. In this paper, we derive efficient, complete, quantum circuits for two representative quantum wavelet transforms, the quantum Haar and quantum Daubechies $D^{(4)}$ transforms. Our approach is to factor the classical operators for these transforms into direct sums, direct products and dot products of unitary matrices. In so doing, we find that permutation matrices, a particular class of unitary matrices, play a pivotal role. Surprisingly, we find that operations that are easy and inexpensive to implement classically are not always easy and inexpensive to implement quantum mechanically, and vice versa. In particular, the computational cost of performing certain permutation matrices is ignored classically because they can be avoided explicitly. However, quantum mechanically, these permutation operations must be performed explicitly and hence their cost enters into the full complexity measure of the quantum transform. We consider the particular set of permutation matrices arising in quantum wavelet transforms and develop efficient quantum circuits that implement them. This allows us to design efficient, complete quantum circuits for the quantum wavelet transform.

Key Words: Quantum Computing, Quantum Algorithms and Circuits, Wavelet Transforms

1 Introduction

The field of quantum computing has undergone an explosion of activity over the past few years. Several important quantum algorithms are now known. Moreover, prototypical quantum computers have already been built using nuclear magnetic resonance [1, 2] and nonlinear optics technologies [3]. Such devices are far from

being general-purpose computers. Nevertheless, they constitute significant milestones along the road to practical quantum computing.

A quantum computer is a physical device whose natural evolution over time can be interpreted as the execution of a useful computation. The basic element of a quantum computer is the quantum bit or "qubit", implemented physically as the state of some convenient 2-state quantum system such as the spin of an electron. Whereas a classical bit must be either a 0 or a 1 at any instant, a qubit is allowed to be an arbitrary superposition of a 0 and a 1 simultaneously. To make a quantum memory register we simply consider the simultaneous state of (possibly entangled) tuples of qubits.

The state of a quantum memory register, or any other isolated quantum system, evolves in time according to some unitary operator. Hence, if the evolved state of a quantum memory register is interpreted as having implemented some computation, that computation must be describable as a unitary operator. If the quantum memory register consists of n qubits, this operator will be represented, mathematically, as some $2^n \times 2^n$ dimensional unitary matrix.

Several quantum algorithms are now known, the most famous examples being Deutsch and Jozsa's algorithm for deciding whether a function is even or balanced [4], Shor's algorithm for factoring a composite integer [5] and Grover's algorithm for finding an item in an unstructured database [6]. However, the field is growing rapidly and new quantum algorithms are being discovered every year. Some recent examples include Brassard, Hoyer, and Tapp's quantum algorithm for counting the number of solutions to a problem [7], Cerf, Grover and Williams quantum algorithm for solving NP-complete problems by nesting one quantum search within another [8] and van Dam, Hoyer, and Tapp's algorithm for distributed quantum computing [9].

The fact that quantum algorithms are describable in terms of unitary transformations is both good news and bad for quantum computing. The good news is that knowing that a quantum computer must perform a unitary transformation allows theorems to be proved about the tasks that quantum computers can and cannot do. For example, Zalka has proved that Grover's algorithm is optimal [10]. Aharonov, Kitaev, and Nisan have proved that a quantum algorithm that involves intermediate measurements is no more powerful than one that postpones all measurements until the end of the unitary evolution stage [11]. Both these proofs rely upon quantum algorithms being unitary transformations. On the other hand, the bad news is that many computations that we would like to perform are not originally described in terms of unitary operators. For example, a desired computation might be nonlinear, irreversible or both nonlinear and irreversible. As a unitary transformation must be linear and reversible we might need to be quite creative in encoding a desired computation on a quantum computer. Irreversibility can be handled by incorporating extra "ancilla" qubits that permit us to remember the input corresponding to each output. But nonlinear transformations are still problematic.

Fortunately, there is an important class of computations, the unitary transforms, such as the Fourier transform, Walsh-Hadamard transform and assorted

wavelet transforms, that are describable, naturally, in terms of unitary operators. Of these, the Fourier and Walsh-Hadamard transforms have been the ones studied most extensively by the quantum computing community. In fact, the quantum Fourier transform (QFT) is now recognized as being pivotal in many known quantum algorithms [12]. The quantum Walsh-Hadamard transform is a critical component of both Shor's algorithm [5] and Grover's algorithm [6]. However, the wavelet transforms are every bit as useful as the Fourier transform, at least in the context of classical computing. For example, wavelet transforms are particularly suited to exposing the multi-scale structure of a signal. They are likely to be useful for quantum image processing and quantum data compression. It is natural therefore to consider how to achieve a quantum wavelet transform.

Starting with the unitary operator for the wavelet transform, the next step in the process of finding a quantum circuit that implements it, is to factor the wavelet operator into the direct sum, direct product and dot product of smaller unitary operators. These operators correspond to 1-qubit and 2-qubit quantum gates. For such a circuit to be physically realizable, the number of gates within it must be bounded above by a polynomial in the number of qubits, n . Finding such a factorization can be extremely challenging. For example, although there are known algebraic techniques for factoring an arbitrary $2^n \times 2^n$ operator, e.g. [13], they are guaranteed to produce $O(2^n)$, i.e., exponentially many, terms in the factorization. Hence, although such a factorization is mathematically valid, it is physically unrealizable because, when treated as a quantum circuit design, would require too many quantum gates. Indeed, Knill has *proved* that an arbitrary unitary matrix will require exponentially many quantum gates if we restrict ourselves to using only gates that correspond to all 1-qubit rotations and XOR [14]. It is therefore clear that the key enabling factor for achieving an efficient quantum implementation, i.e., with a polynomial time and space complexity, is to exploit the specific structure of the given unitary operator.

Perhaps the most striking example of the potential for achieving compact and efficient quantum circuits is the case of the Walsh-Hadamard transform. In quantum computing, this transform arises whenever a quantum register is loaded with all integers in the range 0 to $2^n - 1$. Classically, application of the Walsh-Hadamard transform on a vector of length 2^n involves a complexity of $O(2^n)$. Yet, by exploiting the factorization of the Walsh-Hadamard operator in terms of the Kroenecker product, it can be implemented with a complexity of $O(1)$ by n identical 1-qubit quantum gates. Likewise, the classical FFT algorithm has been found to be implementable in a polynomial space and time complexity, quantum circuit [15] (see also Sec. 2.3). However, exploitation of the operator structure arising in the wavelet transforms (and perhaps other unitary transforms) is more challenging.

A key technique, in classical computing, for exposing and exploiting specific structure of a given unitary transform is the use of permutation matrices. In fact, there is an extensive literature in classical computing on the use of permutation matrices for factorizing unitary transforms into simpler forms that enable

efficient implementations to be devised (see, for example, [16] and [17]). However, the underlying assumption in using the permutation matrices in classical computation is that they can be implemented easily and inexpensively. Indeed, they are considered so trivial that the cost of their implementation is often not included in the complexity analysis. This is because any permutation matrix can be described by its effect on the ordering of the elements of a vector. Hence, it can simply be implemented by re-ordering the elements of the vector involving only data movement and without performing any arithmetic operations. As is shown in this paper, the permutation matrices also play a pivotal role in the factorization of the unitary operators that arise in the wavelet transforms. However, unlike the classical computing, the cost of implementation of the permutation matrices cannot be neglected in quantum computing. Indeed, the main issue in deriving feasible and efficient quantum circuits for the quantum wavelet transforms considered in this paper, is the design of efficient quantum circuits for certain permutation matrices. Note that, any permutation matrix acting on n qubits can mathematically be represented by a $2^n \times 2^n$ unitary operator. Hence, it is possible to factor any permutation matrix by using general techniques such as [13] but this would lead to an exponential time and space complexity. However, the permutation matrices, due to their specific structure (i.e., sparsity pattern), represents a very special subclass of unitary matrices. Therefore, the key to achieve an efficient quantum implementation of permutation matrices is the exploitation of this specific structure.

In this paper, we first develop efficient quantum circuits for a set of permutation matrices arising in the development of the quantum wavelet transforms (and the quantum Fourier transform). We propose three techniques for an efficient quantum implementation of permutation matrices, depending on the permutation matrix considered. In the first technique, we show that a certain class of permutation matrices, designated as *qubit permutation matrices*, can directly be described by their effect on the ordering of qubits. This quantum description is very similar to classical description of the permutation matrices. We show that the *Perfect Shuffle* permutation matrix, designated as Π_{2^n} , and the *Bit Reversal* permutation matrix, designated as P_{2^n} , which arise in the quantum wavelet and Fourier transforms (as well as in many other classical computations) belong to this class. We present a new gate, designated as the *qubit swap gate* or Π_4 , which can be used to directly derive efficient quantum circuits for implementation of the qubit permutation matrices. Interestingly, such circuits for quantum implementation of Π_{2^n} and P_{2^n} lead to new factorizations of these two permutation matrices which were not previously known in classical computation. A second technique is based on a *quantum arithmetic description* of permutation matrices. In particular, we consider the *downshift* permutation matrix, designated as Q_{2^n} , which plays a major role in derivation of quantum wavelet transforms and also frequently arises in many classical computations [16]. We show that a quantum description of Q_{2^n} can be given as a *quantum arithmetic operator*. This description then allows the quantum implementation of Q_{2^n} by using the quantum arithmetic circuits proposed in [18].

A third technique is based on developing totally new factorizations of the permutation matrices. This technique is the most case dependent, challenging, and even counterintuitive (from a classical computing point of view). For this technique, we again consider the permutation matrix Q_{2^n} and we show that it can be factored in terms of FFT which then allows its implementation by using the circuits for QFT. More interestingly, however, we derive a recursive factorization of Q_{2^n} which was not previously known in classical computation. This new factorization enables a direct and efficient implementation of Q_{2^n} . Our analysis of though a limited set of permutation matrices reveals some of the surprises of quantum computing in contrast to classical computing. That is, certain operations that are hard to implement in classical computing are much easier to implement on quantum computing and vice versa. As a specific example, while the classical implementation of Π_{2^n} and P_{2^n} are much harder (in terms of the data movement pattern) than Q_{2^n} , their quantum implementation is much easier and more straightforward than Q_{2^n} .

Given a wavelet kernel, its application is usually performed according to the packet or pyramid algorithms. Efficient quantum implementation of these two algorithms requires efficient circuits for operators of the form $I_{2^{n-i}} \otimes \Pi_{2^i}$ and $\Pi_{2^i} \oplus I_{2^{n-2^i}}$, for some i , where \otimes and \oplus designate, respectively, the kronecker product and the direct sum operator. We show that these operators can be efficiently implemented by using our proposed circuits for implementation of Π_{2^i} . We then consider two representative wavelet kernels, the Haar [17] and Daubechies $D^{(4)}$ [19] wavelets which have previously been considered by Hoyer [20]. For the Haar wavelet, we show that Hoyer's proposed solution is incomplete since it does not lead to a gate-level circuit and, consequently, it does not allow the analysis of time and space complexity. We propose a scheme for design of a complete gate-level circuit for the Haar wavelet and analyze its time and space complexity. For the Daubechies $D^{(4)}$ wavelet, we develop three new factorizations which lead to three gate-level circuits for its implementation. Interestingly, one of this factorization allows efficient implementation of Daubechies $D^{(4)}$ wavelets by using the circuit for QFT.

2 Efficient Quantum Circuits for Two Fundamental Qubits Permutation Matrices: Perfect Shuffle and Bit-Reversal

In this section, we develop quantum circuits for two fundamental permutation matrices, the perfect shuffle, Π_{2^n} , and the bit reversal, P_{2^n} , permutation matrices, which arise in quantum wavelet and Fourier transforms as well as many classical computations involving unitary transforms for signal and image processing [16]. For quantum computing, these two permutation matrices can directly be described in terms of their effect on ordering of qubits. This enables the design of efficient circuits for their implementation. Interestingly, these circuits lead to the discovery of new factorizations for these two permutation matrices.

2.1 Perfect Shuffle Permutation Matrices

A classical description of Π_{2^n} can be given by describing its effect on a given vector. If Z is a 2^n -dimensional vector, then the vector $Y = \Pi_{2^n} Z$ is obtained by splitting Z in half and then shuffling the top and bottom halves of the deck. Alternatively, a description of the matrix Π_{2^n} , in terms of its elements Π_{ij} , for i and $j = 0, 1, \dots, 2^n - 1$, can be given as

$$\Pi_{ij} = \begin{cases} 1 & \text{if } j = i/2 \text{ and } i \text{ is even, or if } j = (i-1)/2 + 2^{n-1} \text{ and } i \text{ is odd} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

As first noted by Hoyer [20], a quantum description of Π_{2^n} can be given by

$$\Pi_{2^n} : |a_{n-1} a_{n-2} \cdots a_1 a_0\rangle \longmapsto |a_0 a_{n-1} a_{n-2} \cdots a_1\rangle \quad (2)$$

That is, for quantum computation, Π_{2^n} is the operator which performs the left qubit-shift operation on n qubits. Note that, $\Pi_{2^n}^t$ (t indicates the transpose) performs the right qubit-shift operation, i.e.,

$$\Pi_{2^n}^t : |a_{n-1} a_{n-2} \cdots a_1 a_0\rangle \longmapsto |a_{n-2} \cdots a_1 a_0 a_{n-1}\rangle \quad (3)$$

2.2 Bit-Reversal Permutation Matrices

A classical description of P_{2^n} can be given by describing its effect on a given vector. If Z is a 2^n -dimensional vector and $Y = P_{2^n} Z$, then $Y_i = Z_j$, for $i = 0, 1, \dots, 2^n - 1$, wherein j is obtained by reversing the bits in the binary representation of index i . Therefore, a description of the matrix P_{2^n} , in terms of its elements P_{ij} , for i and $j = 0, 1, \dots, 2^n - 1$, is given as

$$P_{ij} = \begin{cases} 1 & \text{if } j \text{ is bit reversal of } i \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

A factorization of P_{2^n} in terms of Π_{2^i} is given as [16]

$$P_{2^n} = \Pi_{2^n}(I_2 \otimes \Pi_{2^{n-1}}) \cdots (I_2 \otimes \Pi_{2^{n-i}}) \cdots (I_{2^{n-3}} \otimes \Pi_8)(I_{2^{n-2}} \otimes \Pi_4) \quad (5)$$

A quantum description of P_{2^n} is given as

$$P_{2^n} : |a_{n-1} a_{n-2} \cdots a_1 a_0\rangle \longmapsto |a_0 a_1 \cdots a_{n-2} a_{n-1}\rangle \quad (6)$$

That is, P_{2^n} is the operator which reverses the order of n qubits. This quantum description can be seen from the factorization of P_{2^n} , given by (5), and quantum description of permutation matrices Π_{2^i} . It is interesting to note that for classical computation the term "bit-reversal" refers to reversing the bits in the binary representation of index of the elements of a vector while, for quantum computation, the matrix P_{2^n} literally performs a reversal of the order of qubits.

Note that, P_{2^n} is symmetric, i.e., $P_{2^n} = P_{2^n}^t$ [16]. This can be also easily proved based on the quantum description of P_{2^n} since if the qubits are reversed twice then the original ordering of the qubits is restored. This implies that, $P_{2^n} P_{2^n} = I_{2^n}$ and since P_{2^n} is orthogonal, i.e., $P_{2^n} P_{2^n}^t = I_{2^n}$, it then follows that $P_{2^n} = P_{2^n}^t$.

2.3 Quantum FFT and Bit-Reversal Permutation Matrix

Here, we review the quantum FFT algorithm since it not only arises in derivation of the quantum wavelet transforms (see Sec. 4.3) but also it represents a case in which the roles of permutation matrices Π_{2^n} and P_{2^n} seems to have been overlooked in quantum computing literature.

The classical Cooley-Tukey FFT factorization for a 2^n -dimensional vector is given by [16]

$$F_{2^n} = A_n A_{n-1} \cdots A_1 P_{2^n} = \underline{F}_{2^n} P_{2^n} \quad (7)$$

where $A_i = I_{2^{n-i}} \otimes B_{2^i}$, $B_{2^i} = \frac{1}{\sqrt{2}} \begin{pmatrix} I_{2^{i-1}} & \Omega_{2^{i-1}} \\ I_{2^{i-1}} & -\Omega_{2^{i-1}} \end{pmatrix}$ and $\Omega_{2^{i-1}}$ is a diagonal matrix given as $\Omega_{2^{i-1}} = \text{Diag}\{1, \omega_{2^i}, \omega_{2^i}^2, \dots, \omega_{2^i}^{2^{i-1}-1}\}$ with $\omega_{2^i} = e^{\frac{-2\iota\pi}{2^i}}$ and $\iota = \sqrt{-1}$. We have that $F_2 = W = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. The operator

$$\underline{F}_{2^n} = A_n A_{n-1} \cdots A_1 \quad (8)$$

represents the computational kernel of Cooley-Tukey FFT while P_{2^n} represents the permutation which needs to be performed on the elements of the input vector before feeding that vector into the computational kernel. Note that, the presence of P_{2^n} in (7) is due to the accumulation of its factors, i.e., the terms $(I_{2^i} \otimes \Pi_{2^{n-i}})$, as given by (5).

The Gentleman-Sande FFT factorization is obtained by exploiting the symmetry of F_{2^n} and transposing the Cooley-Tukey factorization [16] leading to

$$F_{2^n} = P_{2^n} A_1^t \cdots A_{n-1}^t A_n^t = P_{2^n} \underline{F}_{2^n}^t \quad (9)$$

where

$$\underline{F}_{2^n}^t = A_1^t \cdots A_{n-1}^t A_n^t \quad (10)$$

represents the computational kernel of the Gentleman-Sande FFT while P_{2^n} represents the permutation which needs to be performed to obtain the elements of the output vector in the correct order.

In [15] a quantum circuit for the implementation of \underline{F}_{2^n} , given by (8), is presented by developing a factorization of the operators B_{2^i} as

$$B_{2^i} = \frac{1}{\sqrt{2}} \begin{pmatrix} I_{2^{i-1}} & \Omega_{2^{i-1}} \\ I_{2^{i-1}} & -\Omega_{2^{i-1}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} I_{2^{i-1}} & I_{2^{i-1}} \\ I_{2^{i-1}} & -I_{2^{i-1}} \end{pmatrix} \begin{pmatrix} I_{2^{i-1}} & 0 \\ 0 & \Omega_{2^{i-1}} \end{pmatrix} \quad (11)$$

Let $C_{2^i} = \begin{pmatrix} I_{2^{i-1}} & 0 \\ 0 & \Omega_{2^{i-1}} \end{pmatrix}$. It then follows that

$$B_{2^i} = (W \otimes I_{2^{i-1}}) C_{2^i} \quad (12)$$

$$A_i = I_{2^{n-i}} \otimes B_{2^i} = (I_{2^{n-i}} \otimes W \otimes I_{2^{i-1}})(I_{2^{n-i}} \otimes C_{2^i}) \quad (13)$$

In [15] a factorization of the operators C_{2^i} is developed as

$$C_{2^i} = \theta_{n-1, n-i} \theta_{n-2, n-i} \cdots \theta_{n-i+1, n-i} \quad (14)$$

where θ_{jk} is a two-bit gate acting on j th and k th qubits.

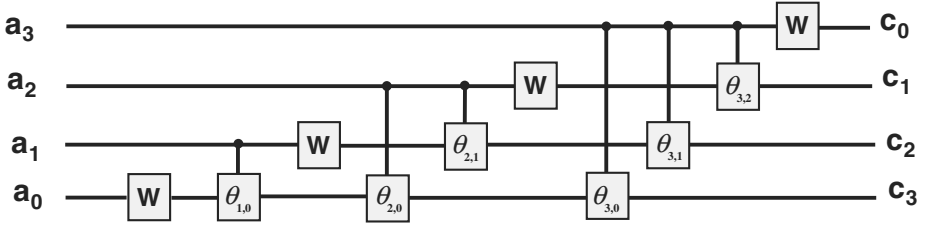


Fig. 1. A circuit for implementation of quantum Fourier transform, QFT (from [15]).

Using (13)-(14) a circuit for implementation of (8) is developed in [15] and presented in Fig. 1. However, there is an error in the corresponding figure in [15] since it implies that, with a correct ordering of the input qubits, the output qubits are obtained in a reverse order. Note that, as can be seen from (7), the operator \underline{F}_{2^n} performs the FFT operation and provides the output qubits in a correct order if the input qubits are presented in a reverse order.

The quantum circuit for Gentleman-Sande FFT can be obtained from the circuit of Fig. 1 by first reversing the order of gates that build the operator block A_i (and thus building operators A_i^t) and then reversing the order of the blocks representing operators A_i . By using the Gentleman-Sande circuit, with the input qubits in the correct order the output qubits are obtained in reverse order.

For an efficient and correct implementation of the quantum FFT, one needs to take into account the ordering of the input and output qubits, particularly if the FFT is used as a block box in a quantum computation. If the FFT is used as a stand-alone block or as the last stage in the computation (and hence its output is sampled directly), then it is more efficient to use the Gentleman-Sande FFT since the ordering of the output qubits does not cause any problem. If the FFT is used as the first stage of the computation, then it is more efficient to use the Cooley-Tukey factorization by preparing the input qubits in a reverse order. Note that, as in classical computation, each or a combination of the Cooley-Tukey or Gentleman-Sande FFT factorization can be chosen in a given quantum computation to avoid explicit implementation of P_{2^n} (or, any other mechanism) for reversing the order of qubits and hence achieve a greater efficiency. As an example, in Sec. 4.3 we will show that the use of the Cooley-Tukey rather than the Gentleman-Sande factorization leads to a greater efficiency in quantum implementation by eliminating the need for an explicit implementation of P_{2^n} (or, any other mechanism) for reversing the order of qubits.

2.4 A Basic Quantum Gate for Efficient Implementation of Qubits Permutation Matrices

If a permutation matrix can be described by its effect on the ordering of the qubits then it might be possible to devise circuits for its implementation directly. We call the class of such permutation matrices as "Qubit Permutation Matrices".

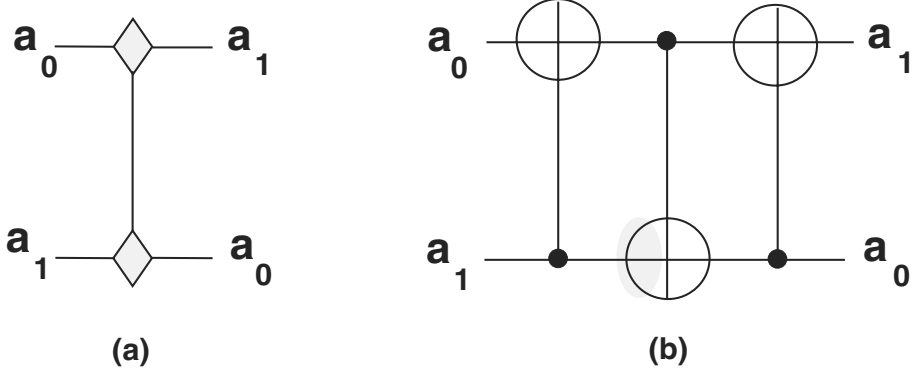


Fig. 2. The Π_4 gate (a) and its implementation by using three XOR (Controlled-NOT) gates (b).

A set of efficient and practically realizable circuits for implementation of Qubit Permutation Matrices can be built by using a new quantum gate, called *the qubit swap gate*, Π_4 , where

$$\Pi_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (15)$$

For quantum computation, Π_4 is the "qubit swap operator", i.e.,

$$\Pi_4 : |a_1 a_0\rangle \mapsto |a_0 a_1\rangle \quad (16)$$

The Π_4 gate, shown in Fig. 2.a, can be implemented with three XOR (or Controlled-NOT) gates as shown in Fig. 2.b. The Π_4 gate offers two major advantages for practical implementation:

- It performs a local operation, i.e., swapping the two neighboring qubits. This locality can be advantageous in practical realizations of quantum circuits, and
- Given the fact that Π_4 can be implemented using three XOR (or, Controlled-NOT) gates, it is possible to implement conditional operators involving Π_4 , for example, operators of the form $\Pi_4 \oplus I_{2^{n-4}}$, by using Controlled^k-NOT gates [21].

A circuit for implementation of Π_{2^n} by using Π_4 gates is shown in Fig. 3. This circuit is based on an intuitively simple idea of successive swapping of the neighboring qubits, and implements Π_{2^n} with a complexity of $O(n)$ by using an $O(n)$ number of Π_4 gates. It is interesting to note that, this circuit leads to a new (to our knowledge) factorization of Π_{2^n} in terms of Π_4 as

$$\begin{aligned} \Pi_{2^n} = & (I_{2^{n-2}} \otimes \Pi_4)(I_{2^{n-3}} \otimes \Pi_4 \otimes I_2) \cdots (I_{2^{n-i}} \otimes \Pi_4 \otimes I_{2^{i-2}}) \times \\ & \cdots (I_2 \otimes \Pi_4 \otimes I_{2^{n-3}})(\Pi_4 \otimes I_{2^{n-2}}) \end{aligned} \quad (17)$$

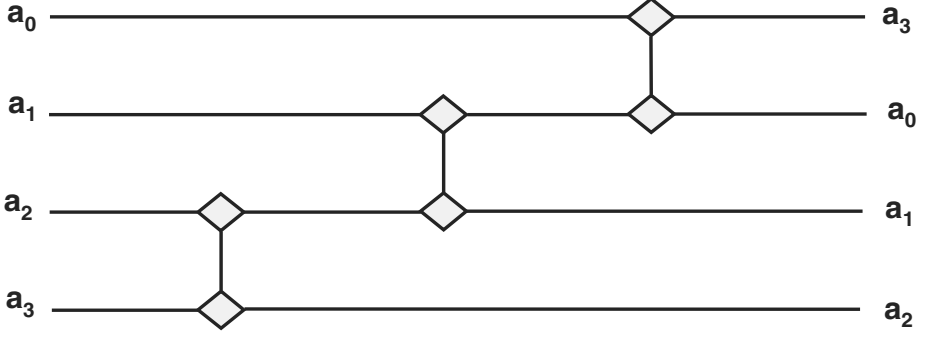


Fig. 3. A circuit for implementation of Perfect Shuffle permutation matrix, Π_{2^n} .

This new factorization of Π_{2^n} is less efficient than other schemes (see, for example, [16]) for a *classical implementation* of Π_{2^n} . Interestingly, it is derived here as a result of our search for an efficient *quantum implementation* of Π_{2^n} , and in this sense it is only efficient for a quantum implementation. Note also, that a new (to our knowledge) recursive factorization of Π_{2^i} directly results from Fig. (3) as

$$\Pi_{2^i} = (I_{2^{i-2}} \otimes \Pi_4)(\Pi_{2^{i-1}} \otimes I_2) \quad (18)$$

A circuit for implementation of P_{2^n} by using Π_4 gates is shown in Fig. 4. Again, this circuit is based on an intuitively simple idea, that is, successive and parallel swapping of the neighboring qubits, and implements P_{2^n} with a complexity of $O(n)$ by using $O(n^2)$ Π_4 gates. This circuit leads to a new (to our knowledge) factorization of P_{2^n} in terms of Π_4 as

$$P_{2^n} = \left(\underbrace{(\Pi_4 \otimes \Pi_4 \cdots \otimes \Pi_4)}_{\frac{n}{2}} (I_2 \otimes \underbrace{\Pi_4 \otimes \cdots \otimes \Pi_4}_{\frac{n}{2}-1} \otimes I_2) \right)^{\frac{n}{2}} \quad (19)$$

for n even, and

$$P_{2^n} = \left((I_2 \otimes \underbrace{\Pi_4 \otimes \cdots \otimes \Pi_4}_{\frac{n-1}{2}}) \underbrace{(\Pi_4 \otimes \cdots \otimes \Pi_4)}_{\frac{n-1}{2}} \otimes I_2 \right)^{\frac{n-1}{2}} (I_2 \otimes \underbrace{\Pi_4 \otimes \cdots \otimes \Pi_4}_{\frac{n-1}{2}}) \quad (20)$$

for n odd.

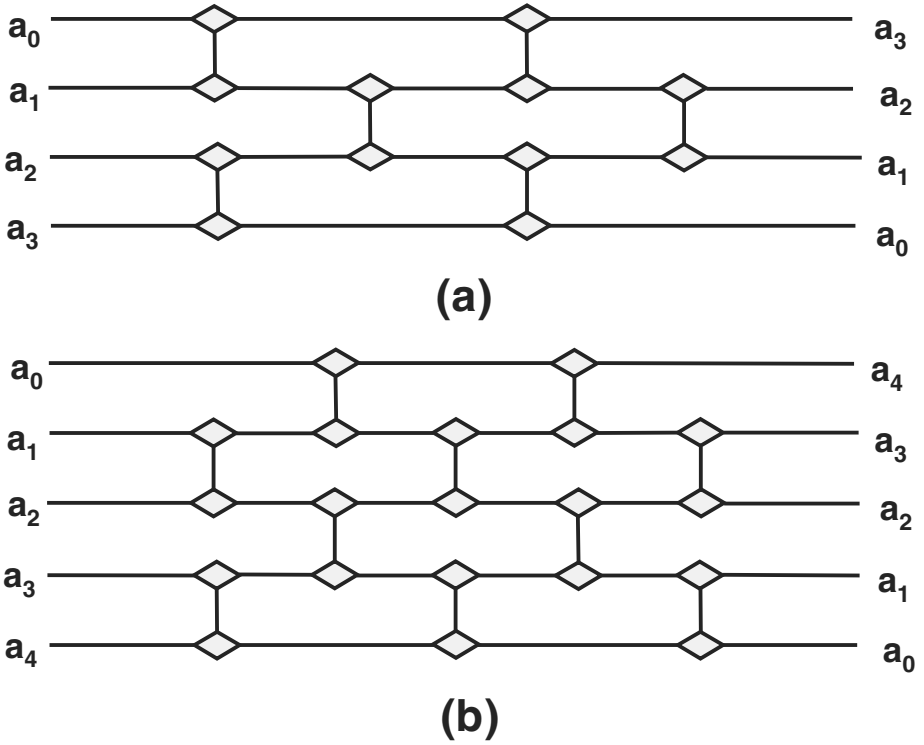


Fig. 4. Circuits for implementation of Bit Reversal permutation matrix, P_{2^n} , for n even (a) and for n odd (b).

It should be emphasized that this new factorization of P_{2^n} is less efficient than other schemes, e.g., the use of (5) for a *classical implementation* (see also [16] for further discussion). However, this factorization is more efficient for a *quantum implementation* of P_{2^n} . In fact, a quantum implementation of P_{2^n} by using (5) and (17) will result in a complexity of $O(n^2)$ by using $O(n^2)$ Π_4 gates.

As will be shown, the development of *complete* and efficient circuits for implementation of wavelet transforms requires a mechanism for implementation of conditional operators of the forms $\Pi_{2^i} \oplus I_{2^n-2^i}$ and $P_{2^i} \oplus I_{2^n-2^i}$, for some i . The key enabling factor for a successful implementation of such conditional operators is the use of factorizations similar to (17) and (19)-(20) or, alternatively, circuits similar to those in Figures 3 and 4, along with the conditional operators involving Π_4 gates.

3 Quantum Wavelet Algorithms

3.1 Wavelet Pyramidal and Packet Algorithms

Given a wavelet kernel, its corresponding wavelet transform is usually performed according to a packet algorithm (PAA) or a pyramid algorithm (PYA). The first step in devising quantum counterparts of these algorithms is the development of suitable factorizations. Consider the Daubechies fourth-order wavelet kernel of dimension 2^i , denoted as $D_{2^i}^{(4)}$. First level factorizations of PAA and PYA for a 2^n -dimensional vector are given as

$$PAA = (I_{2^{n-2}} \otimes D_4^{(4)})(I_{2^{n-3}} \otimes \Pi_8) \cdots (I_{2^{n-i}} \otimes D_{2^i}^{(4)}) \times \\ (I_{2^{n-i-1}} \otimes \Pi_{2^{i+1}}) \cdots (I_2 \otimes D_{2^{n-1}}^{(4)}) \Pi_{2^n} D_{2^n}^{(4)} \quad (21)$$

$$PYA = (D_4^{(4)} \oplus I_{2^{n-4}})(\Pi_8 \oplus I_{2^{n-8}}) \cdots (D_{2^i}^{(4)} \oplus I_{2^{n-2^i}}) \\ (\Pi_{2^{i+1}} \oplus I_{2^{n-2^{i+1}}}) \cdots \Pi_{2^n} D_{2^n}^{(4)} \quad (22)$$

These factorizations allow a first level analysis of the feasibility and efficiency of quantum implementations of the packet and pyramid algorithms. To see this, suppose we have a practically realizable and efficient, i.e., $O(i)$, quantum algorithm for implementation of $D_{2^i}^{(4)}$. For the packet algorithm, the operators $(I_{2^{n-i}} \otimes D_{2^i}^{(4)})$ can be directly and efficiently implemented by using the algorithm for $D_{2^i}^{(4)}$. Also, using the factorization of Π_{2^i} , given by (17), the operators $(I_{2^{n-i}} \otimes \Pi_{2^i})$ can be implemented efficiently in $O(i)$.

For the pyramid algorithm, the existence of an algorithm for $D_{2^i}^{(4)}$ does not automatically imply an efficient algorithm for implementation of the conditional operators $(D_{2^i}^{(4)} \oplus I_{2^{n-2^i}})$. An example of such a case is discussed in Sec. 4.4. Thus, careful analysis is needed to establish both the feasibility and efficiency of implementation of the conditional operators $(D_{2^i}^{(4)} \oplus I_{2^{n-2^i}})$ by using the algorithm for $D_{2^i}^{(4)}$. Note, however, that the conditional operators $(\Pi_{2^i} \oplus I_{2^{n-2^i}})$ can be efficiently implemented in $O(i)$ by using the factorization in (17) and the conditional Π_4 gates.

The above analysis can be extended to any wavelet kernel (WK) and summarized as follows:

- Packet algorithm: A physically realizable and efficient algorithm for the WK along with the use of (17) leads to a physically realizable and efficient implementation of the packet algorithm.
- Pyramid algorithm: A physically realizable and efficient algorithm for the WK does not automatically lead to an implementation of the conditional operators involving WK (and hence the pyramid algorithm) but the conditional operators $(\Pi_{2^i} \oplus I_{2^{n-2^i}})$ can be efficiently implemented by using the factorization in (17) and the conditional Π_4 gates.

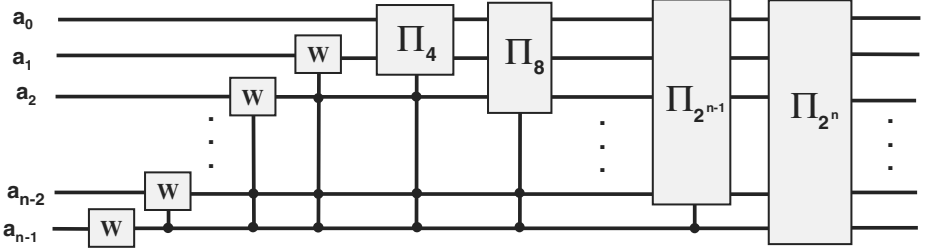


Fig. 5. A block-level circuit for Haar wavelet (from [20]).

3.2 Haar Wavelet Factorization and Implementation

The Haar transform can be defined from the Haar functions [17]. Hoyer [20] used a recursive definition of Haar matrices based on the *generalized Kronecker product* (see also [17] for similar definitions) and developed a factorization of H_{2^n} as

$$H_{2^n} = (I_{2^{n-1}} \otimes W) \cdots (I_{2^{n-i}} \otimes W \oplus I_{2^{n-2^{n-i}+1}}) \cdots (W \oplus I_{2^n-2}) \times \\ (\Pi_4 \oplus I_{2^n-4}) \cdots (\Pi_{2^i} \oplus I_{2^n-2^i}) \cdots (\Pi_{2^{n-1}} \oplus I_{2^n-1}) \Pi_{2^n} \quad (23)$$

Hoyer's circuit for implementation of (23) is shown in Fig 5. However, this represents an *incomplete* solution for quantum implementation and subsequent complexity analysis of the Haar transform. To see this, let

$$H_{2^n}^{(1)} = (I_{2^{n-1}} \otimes W) \cdots (I_{2^{n-i}} \otimes W \oplus I_{2^{n-2^{n-i}+1}}) \cdots (W \oplus I_{2^n-2}) \quad (24)$$

$$H_{2^n}^{(2)} = (\Pi_4 \oplus I_{2^n-4}) \cdots (\Pi_{2^i} \oplus I_{2^n-2^i}) \cdots (\Pi_{2^{n-1}} \oplus I_{2^n-1}) \Pi_{2^n} \quad (25)$$

Clearly, the operator $H_{2^n}^{(1)}$ can be implemented in $O(n)$ by using $O(n)$ conditional W gates. But the feasibility of practical implementation of the operator $H_{2^n}^{(2)}$ and its complexity (and consequently those of the factorization in (23)) cannot be assessed unless a mechanism for implementation of the terms $(\Pi_{2^i} \oplus I_{2^n-2^i})$ is devised.

However, by using the factorizations and circuits similar to (17) and Figure 3, it can be easily shown that the operators $(\Pi_{2^i} \oplus I_{2^n-2^i})$ can be implemented in $O(i)$ by using $O(i)$ conditional Π_4 gates (or, Controlled^k-NOT gates). This leads to the implementation of $H_{2^n}^{(2)}$ and consequently H_{2^n} in $O(n^2)$ by using $O(n^2)$ gates. This represents not only the first practically feasible quantum circuit for implementation of H_{2^n} but also the first complete analysis of complexity of its time and space (gates) quantum implementation. Note that, both operators $(I_{2^{n-i}} \otimes H_{2^i})$ and $(H_{2^i} \oplus I_{2^n-2^i})$ can be directly and efficiently implemented by using the above algorithm and circuit for implementation of H_{2^i} . This implies both the feasibility and efficiency of the quantum implementation of the packet and pyramid algorithms by using our factorization for Haar wavelet kernel.

3.3 Daubechies $D^{(4)}$ Wavelet and Hoyer's Factorization

The Daubechies fourth-order wavelet kernel of dimension 2^n is given in a matrix form as [22]

$$D_{2^n}^{(4)} = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & & & & \\ c_3 & -c_2 & c_1 & -c_0 & & & & \\ & & c_0 & c_1 & c_2 & c_3 & & \\ & & c_3 & -c_2 & c_1 & -c_0 & & \\ \vdots & \vdots & & & & & \ddots & \\ & & & & & c_0 & c_1 & c_2 & c_3 \\ & & & & & c_3 & -c_2 & c_1 & -c_0 \\ c_2 & c_3 & & & & & & c_0 & c_1 \\ c_1 & -c_0 & & & & & & c_3 & -c_2 \end{pmatrix} \quad (26)$$

where $c_0 = \frac{(1+\sqrt{3})}{4\sqrt{2}}$, $c_1 = \frac{(3+\sqrt{3})}{4\sqrt{2}}$, $c_2 = \frac{(3-\sqrt{3})}{4\sqrt{2}}$, and $c_3 = \frac{(1-\sqrt{3})}{4\sqrt{2}}$. For classical computation and given its sparse structure, the application of $D_{2^n}^{(4)}$ can be performed with an optimal cost of $O(2^n)$. However, the matrix $D_{2^n}^{(4)}$, as given by (26), is not suitable for a quantum implementation. To achieve a feasible and efficient quantum implementation, a suitable factorization of $D_{2^n}^{(4)}$ needs to be developed. Hoyer [20] proposed a factorization of $D_{2^n}^{(4)}$ as

$$D_{2^n}^{(4)} = (I_{2^{n-1}} \otimes C_1) S_{2^n} (I_{2^{n-1}} \otimes C_0) \quad (27)$$

where

$$C_0 = 2 \begin{pmatrix} c_4 & -c_2 \\ -c_2 & c_4 \end{pmatrix} \text{ and } C_1 = \frac{1}{2} \begin{pmatrix} \frac{c_0}{c_4} & 1 \\ 1 & \frac{c_1}{c_2} \end{pmatrix} \quad (28)$$

and S_{2^n} is a permutation matrix with a classical description given by

$$S_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } i \text{ is even, or if } i + 2 = j \pmod{2^n} \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

Hoyer's block-level circuit for implementation of (27) is shown in Figure 6. Clearly, the main issue for a practical quantum implementation and subsequent complexity analysis of (27) is the quantum implementation of matrix S_{2^n} . To this end, Hoyer discovered a quantum arithmetic description of S_{2^n} as

$$S_{2^n} : |a_{n-1} a_{n-2} \cdots a_1 a_0\rangle \mapsto |b_{n-1} b_{n-2} \cdots b_1 b_0\rangle \quad (30)$$

where

$$b_i = \begin{cases} a_i - 2 \pmod{n}, & \text{if } i \text{ is odd} \\ a_i & \text{otherwise} \end{cases} \quad (31)$$

As suggested by Hoyer, this description of S_{2^n} then allows its quantum implementation by using quantum arithmetic circuits of [18] with a complexity of $O(n)$. This algorithm can be directly extended for implementation of the operators $(I_{2^{n-i}} \otimes D_{2^i}^{(4)})$ and hence the packet algorithm. However, the feasibility and efficiency of an implementation of the operators $(I_{2^{n-i}} \oplus D_{2^i}^{(4)})$ and thus the pyramid algorithm needs further analysis.

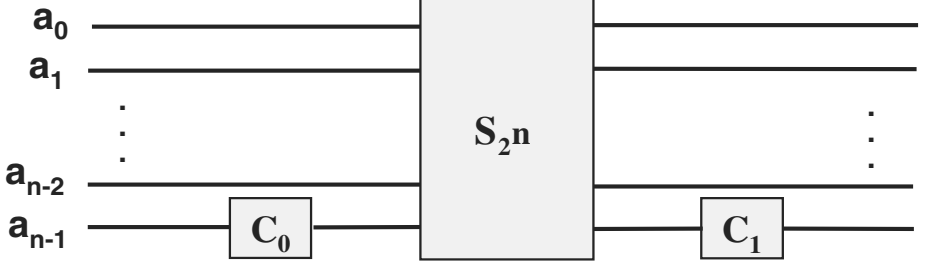


Fig. 6. A block-level circuit for implementation of Hoyer's factorization of $D_{2^n}^{(4)}$.

4 Fast Quantum Algorithms and Circuits for Implementation of Daubechies $D^{(4)}$ Wavelet

In this section, we develop a new factorization of the Daubechies $D^{(4)}$ wavelet. This factorization leads to three new and efficient circuits, including one using the circuit for QFT, for implementation of Daubechies $D^{(4)}$ wavelet.

4.1 A New Factorization of Daubechies $D^{(4)}$ Wavelet

We develop a new factorization of the Daubechies $D^{(4)}$ wavelet transform by showing that the permutation matrix S_{2^n} can be written as a product of two permutation matrices as

$$S_{2^n} = Q_{2^n} R_{2^n} \quad (32)$$

where Q_{2^n} is the *downshift permutation matrix* [16] given by

$$Q_{2^n} = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ 0 & 0 & 0 & 1 & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \quad (33)$$

and R_{2^n} is a permutation matrix given by

$$R_{2^n} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & 0 & 1 \\ & & & 1 & 0 \end{pmatrix} \quad (34)$$

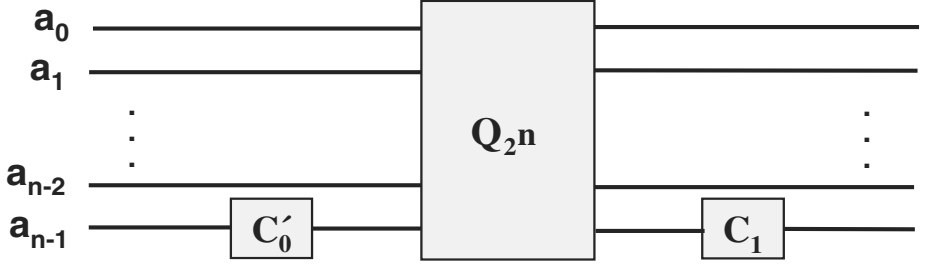


Fig. 7. A block-level circuit for implementation of new factorization of $D_{2^n}^{(4)}$.

The matrix R_{2^n} can be written as

$$R_{2^n} = I_{2^{n-1}} \otimes N \quad (35)$$

where $N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Substituting (35) and (32) into (27), a new factorization of $D_{2^n}^{(4)}$ is derived as

$$D_{2^n}^{(4)} = (I_{2^{n-1}} \otimes C_1) Q_{2^n} (I_{2^{n-1}} \otimes N) (I_{2^{n-1}} \otimes C_0) = (I_{2^{n-1}} \otimes C_1) Q_{2^n} (I_{2^{n-1}} \otimes C'_0) \quad (36)$$

where

$$C'_0 = N.C_0 = 2 \begin{pmatrix} -c_2 & c_4 \\ c_4 & -c_2 \end{pmatrix} \quad (37)$$

Fig. 7 shows a block-level implementation of (36). Clearly, the main issue for a practical quantum gate-level implementation and subsequent complexity analysis of (36) is the quantum implementation of matrix Q_{2^n} . In the following, we present three circuits for quantum implementation of matrix Q_{2^n} .

4.2 Quantum Arithmetic Implementation of Permutation Matrix Q_{2^n}

A first circuit for implementation of matrix Q_{2^n} is developed based on its description as a *quantum arithmetic operator*. We have discovered such a quantum arithmetic description of Q_{2^n} as

$$Q_{2^n} : |a_{n-1} a_{n-2} \cdots a_1 a_0\rangle \mapsto |b_{n-1} b_{n-2} \cdots b_1 b_0\rangle \quad (38)$$

where

$$b_i = a_i - 1 \pmod{n} \quad (39)$$

This description of Q_{2^n} allows its quantum implementation by using quantum arithmetic circuit of [18] with a complexity of $O(n)$. Note, however, that the arithmetic description of Q_{2^n} is simpler than that of S_{2^n} since it does not involve conditional quantum arithmetic operations (i.e., the same operation is applied

to all qubits). This algorithm for quantum implementation of Q_{2^n} and hence $D_{2^n}^{(4)}$ can be directly extended for implementation of the operators $(I_{2^{n-i}} \otimes D_{2^i}^{(4)})$ and hence the packet algorithm. However, the feasibility and efficiency of an implementation of the operators $(I_{2^{n-i}} \oplus D_{2^i}^{(4)})$ and thus the pyramid algorithm needs further analysis.

4.3 Quantum FFT Factorization of Permutation Matrix Q_{2^n}

A direct and efficient factorization and subsequent circuit for implementation of Q_{2^n} (and hence Daubechies $D^{(4)}$ wavelet) can be derived by using the FFT algorithm. This factorization is based on the observation that Q_{2^n} can be described in terms of FFT as [16]

$$Q_{2^n} = F_{2^n} T_{2^n} F_{2^n}^* \quad (40)$$

where T_{2^n} is a diagonal matrix given as $T_{2^n} = \text{Diag}\{1, \omega_{2^n}, \omega_{2^n}^2, \dots, \omega_{2^n}^{2^n-1}\}$ with $\omega_{2^n} = e^{\frac{-2i\pi}{2^n}}$ (* indicates conjugate transpose). As will be seen, it is more efficient to use the Cooley-Tukey factorization, given by (7), and write (40) as

$$Q_{2^n} = \underline{F}_{2^n} P_{2^n} T_{2^n} P_{2^n} \underline{F}_{2^n}^* \quad (41)$$

It can be shown that the matrix T_{2^n} has a factorization as

$$T_{2^n} = (G(\omega_{2^n}^{2^{n-1}}) \otimes I_{2^{n-1}}) \cdots (I_{2^{i-1}} \otimes G(\omega_{2^n}^{2^{n-i}}) \otimes I_{2^{i-1}}) \cdots (I_{2^{n-1}} \otimes G(\omega_{2^n})) \quad (42)$$

where $G(\omega_{2^n}^k) = \text{Diag}\{1, \omega_{2^n}^k\} = \begin{pmatrix} 1 & 0 \\ 0 & \omega_{2^n}^k \end{pmatrix}$. This factorization leads to an efficient implementation of T_{2^n} by using n single qubit $G(\omega_{2^n}^k)$ gates as shown in Fig. 8. Together with the circuit for implementation of P_{2^n} (Fig. 4) and the circuit for implementation of FFT (Fig. 1), they represent a complete gate-level implementation of $D_{2^n}^{(4)}$.

However, a more efficient circuit can be derived by avoiding the explicit implementation of P_{2^n} by showing that the operator

$$P_{2^n} T_{2^n} P_{2^n} = P_{2^n} (G(\omega_{2^n}^{2^{n-1}}) \otimes I_{2^{n-1}}) \cdots (I_{2^{i-1}} \otimes G(\omega_{2^n}^{2^{n-i}}) \otimes I_{2^{i-1}}) \times \cdots (I_{2^{n-1}} \otimes G(\omega_{2^n})) P_{2^n} \quad (43)$$

can be efficiently implemented by simply reversing the order of gates in Fig. 8. This is established by the following lemma:

Lemma 1.

$$P_{2^n} (G(\omega_{2^n}^{2^{n-1}}) \otimes I_{2^{n-1}}) = (I_{2^{n-1}} \otimes G(\omega_{2^n}^{2^{n-1}})) P_{2^n} \quad (44)$$

$$P_{2^n} (I_{2^{n-j}} \otimes G(\omega_{2^n}^{2^{j-1}}) \otimes I_{2^{j-1}}) = (I_{2^{j-1}} \otimes G(\omega_{2^n}^{2^{j-1}}) \otimes I_{2^{n-j}}) P_{2^n} \quad (45)$$

$$P_{2^n} (I_{2^{n-1}} \otimes G(\omega_{2^n})) = (G(\omega_{2^n}) \otimes I_{2^{n-1}}) P_{2^n} \quad (46)$$

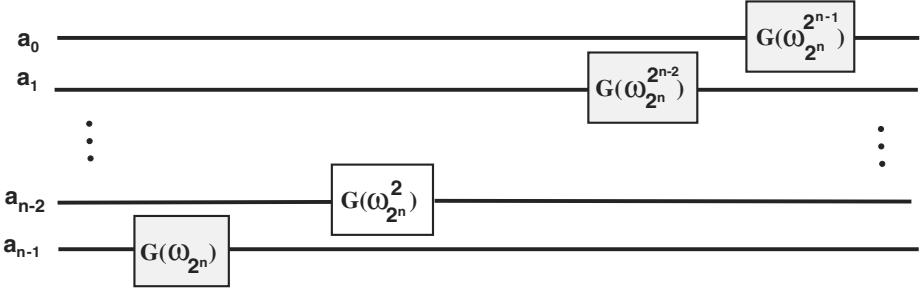


Fig. 8. A circuit for implementation of operator T_{2^n} .

Proof. This lemma can be easily proved based on the physical interpretation of operations in (44)-(46). The left-hand side of (44) implies first an operation, i.e., application of $G(\omega_{2^n}^{2^{n-1}})$, on the last qubit and then application of P_{2^n} on all the qubits, i.e., reversing the order of qubits. However, this is equivalent to first reversing the order of qubits, i.e., applying P_{2^n} , and then applying $G(\omega_{2^n}^{2^{n-1}})$, on the first qubit which is the operation described by the right-hand side of (44). Similarly, the left-hand side of (45) implies first application of $G(\omega_{2^n}^{2^{i-1}})$ on the $(n-i)$ th qubit and then reversing the order of qubits. This is equivalent to first reversing the order of qubits and then applying $G(\omega_{2^n}^{2^{i-1}})$ on the i th qubit which is the operations described by the right hand side of (45). In a same fashion, the left hand side of (46) implies first application of $G(\omega_{2^n}^2)$ on the first qubit and then reversing the order of qubits which is equivalent to first reversing the order of qubits and then applying $G(\omega_{2^n}^2)$ on the last qubit, that is, the operations in right-hand side of (46).

Applying (44)-(46) to (43) from left to right and noting that, due to the symmetry of P_{2^n} , we have $P_{2^n} P_{2^n} = I_{2^n}$, it then follows that

$$P_{2^n} T_{2^n} P_{2^n} = (I_{2^{n-1}} \otimes G(\omega_{2^n}^{2^{n-1}})) \cdots (I_{2^{n-i}} \otimes G(\omega_{2^n}^{2^{n-i}}) \otimes I_{2^{i-1}}) \cdots (G(\omega_{2^n}^2) \otimes I_{2^{n-1}}) \quad (47)$$

The circuit for implementation of (47) is shown in Fig.9 which, as can be seen, has been obtained by reversing the order of gates in Fig. 8. Note that, the use of (47), which is a direct consequence of using the Cooley-Tukey factorization, enables the implementation of (40) without explicit implementation of P_{2^n} .

Using (40) and (47), the complexity of the implementation of Q_{2^n} and thus $D_{2^n}^{(4)}$ is the same as of the quantum FFT, that is, $O(n^2)$ for an exact implementation and $O(nm)$ for an approximation of order m [15]. Note that, by using (47), (40), and (36) both operators $(I_{2^{n-i}} \otimes D_{2^i}^{(4)})$ and $(D_{2^i}^{(4)} \oplus I_{2^{n-2i}})$ can be directly implemented. This implies both the feasibility and efficiency of the quantum implementation of the packet and pyramid algorithms by using this algorithm for quantum implementation of $D_{2^n}^{(4)}$.

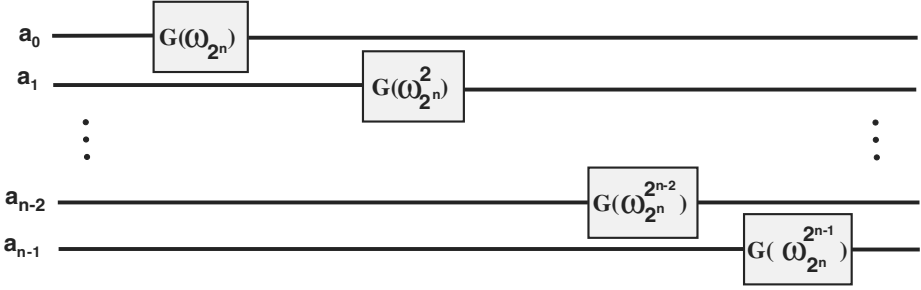


Fig. 9. A circuit for implementation of operator $P_{2^n} T_{2^n} P_{2^n}$

4.4 A Direct Recursive Factorization of Permutation Matrix Q_{2^n}

A new direct and recursive factorization of Q_{2^n} can be derived based on a similarity transformation of Q_{2^n} by using Π_{2^n} as

$$\Pi_{2^n}^t Q_{2^n} \Pi_{2^n} = \begin{pmatrix} 0 & I_{2^{n-1}} \\ Q_{2^{n-1}} & 0 \end{pmatrix} \quad (48)$$

which can be written as

$$\Pi_{2^n}^t Q_{2^n} \Pi_{2^n} = \begin{pmatrix} 0 & I_{2^{n-1}} \\ I_{2^{n-1}} & 0 \end{pmatrix} \begin{pmatrix} Q_{2^{n-1}} & 0 \\ 0 & I_{2^{n-1}} \end{pmatrix} = (N \otimes I_{2^{n-1}})(Q_{2^{n-1}} \oplus I_{2^{n-1}}) \quad (49)$$

from which Q_{2^n} can be calculated as

$$Q_{2^n} = \Pi_{2^n}(N \otimes I_{2^{n-1}})(Q_{2^{n-1}} \oplus I_{2^{n-1}})\Pi_{2^n}^t \quad (50)$$

Replacing a similar factorization of $Q_{2^{n-1}}$ into (50), we get

$$Q_{2^n} = \Pi_{2^n}(N \otimes I_{2^{n-1}})(\Pi_{2^{n-1}}(N \otimes I_{2^{n-2}})(Q_{2^{n-2}} \oplus I_{2^{n-2}})\Pi_{2^{n-1}}^t \oplus I_{2^{n-1}})\Pi_{2^n}^t \quad (51)$$

By using the identity

$$\Pi_{2^{n-1}} A \Pi_{2^{n-1}}^t \oplus I_{2^{n-1}} = (I_2 \otimes \Pi_{2^{n-1}})(A \oplus I_{2^{n-1}})(I_2 \otimes \Pi_{2^{n-1}}^t) \quad (52)$$

for any matrix $A \in \mathbb{R}^{2^{n-1} \times 2^{n-1}}$, (51) can be then written as

$$Q_{2^n} = \Pi_{2^n}(N \otimes I_{2^{n-1}})(I_2 \otimes \Pi_{2^{n-1}})((N \otimes I_{2^{n-2}}) \times (Q_{2^{n-2}} \oplus I_{2^{n-2}}) \oplus I_{2^{n-1}})(I_2 \otimes \Pi_{2^{n-1}}^t)\Pi_{2^n}^t \quad (53)$$

Using the identity

$$\begin{aligned} (N \otimes I_{2^{n-2}})(Q_{2^{n-2}} \oplus I_{2^{n-2}}) \oplus I_{2^{n-1}} &= (N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \\ &\quad (Q_{2^{n-2}} \oplus I_{2^{n-2}} \oplus I_{2^{n-1}}) \\ &= (N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \\ &\quad (Q_{2^{n-2}} \oplus I_{3.2^{n-2}}) \end{aligned} \quad (54)$$

(53) is now written as

$$Q_{2^n} = \Pi_{2^n}(N \otimes I_{2^{n-1}})(I_2 \otimes \Pi_{2^{n-1}})(N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \times \\ (Q_{2^{n-2}} \oplus I_{2^{n-2}n-2})(I_2 \otimes \Pi_{2^{n-1}}^t) \Pi_{2^n}^t \quad (55)$$

Repeating the same procedures for all Q_{2^i} , for $i = n - 3$ to 1, and noting that $Q_2 = N$, it then follows

$$Q_{2^n} = \Pi_{2^n}(N \otimes I_{2^{n-1}})(I_2 \otimes \Pi_{2^{n-1}})(N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \cdots (I_{2^{n-2}} \otimes \Pi_4) \times \\ (N \otimes I_2 \oplus I_{2^{n-4}})(N \otimes I_{2^{n-2}})(I_{2^{n-2}} \otimes \Pi_4^t) \cdots (I_2 \otimes \Pi_{2^{n-1}}^t) \Pi_{2^n}^t \quad (56)$$

The above expression of Q_{2^n} can be further simplified by exploiting the fact that (see Appendix for the proof) every operator of the form $(I_{2^i} \otimes \Pi_{2^{n-i}})$, for $i = n - 2$ to 1, commutes with all operators of the form $(N \otimes I_{2^{n-j}} \oplus I_{2^{n-2}n-j+1})$, for $j = i$ to 1. Using this commutative property, (56) can be now written as

$$Q_{2^n} = \Pi_{2^n}(I_2 \otimes \Pi_{2^{n-1}})(I_4 \otimes \Pi_{2^{n-2}}) \cdots (I_{2^{n-2}} \otimes \Pi_4)(N \otimes I_{2^{n-1}}) \times \\ \cdots (N \otimes I_2 \oplus I_{2^{n-4}})(N \otimes I_{2^{n-2}})(I_{2^{n-2}} \otimes \Pi_4^t) \cdots (I_2 \otimes \Pi_{2^{n-1}}^t) \Pi_{2^n}^t \quad (57)$$

Using the factorization of P_{2^n} given in (5), we then have

$$Q_{2^n} = P_{2^n}(N \otimes I_{2^{n-1}})(N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \cdots (N \otimes I_2 \oplus I_{2^{n-4}})(N \otimes I_{2^{n-2}}) P_{2^n} \quad (58)$$

Substituting (58) into (36), a factorization of $D_{2^n}^{(4)}$ is then obtained as

$$D_{2^n}^{(4)} = (I_{2^{n-1}} \otimes C_1) P_{2^n}(N \otimes I_{2^{n-1}})(N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \cdots (N \otimes I_2 \oplus I_{2^{n-4}}) \times \\ (N \otimes I_{2^{n-2}}) P_{2^n}(I_{2^{n-1}} \otimes C'_0) \quad (59)$$

Using Lemma 1, it then follows that

$$D_{2^n}^{(4)} = P_{2^n}(C_1 \otimes I_{2^{n-1}})(N \otimes I_{2^{n-1}})(N \otimes I_{2^{n-2}} \oplus I_{2^{n-1}}) \cdots (N \otimes I_2 \oplus I_{2^{n-4}}) \times \\ (N \otimes I_{2^{n-2}})(C'_0 \otimes I_{2^{n-1}}) P_{2^n} \quad (60)$$

A circuit for implementation of $D_{2^n}^{(4)}$, based on (60), is shown in Fig. 10. Together with the circuit for implementation of P_{2^n} , shown in Fig. 4, they represent a complete gate-level circuit for implementation of $D_{2^n}^{(4)}$ with an optimal complexity of $O(n)$.

Using (60) and (19)-(20), the operators $(I_{2^{n-i}} \otimes D_{2^i}^{(4)})$ can be directly and efficiently implemented with a complexity of $O(i)$. This implies both the feasibility and efficiency of the implementation of the packet algorithm by using this algorithm for $D_{2^n}^{(4)}$ wavelet kernel. However, this algorithm is less efficient for implementation of the operators $(D_{2^i}^{(4)} \oplus I_{2^{n-2^i}})$ and hence the pyramid algorithm. To see this, note that, the implementation of the operators $(D_{2^i}^{(4)} \oplus I_{2^{n-2^i}})$, by using (60), requires the implementation of the conditional operators $(P_{2^i} \oplus I_{2^{n-2^i}})$. However, these conditional operators cannot be directly implemented by using (19) and (20). An alternative solution is to use the factorization of P_{2^i} in (5) and

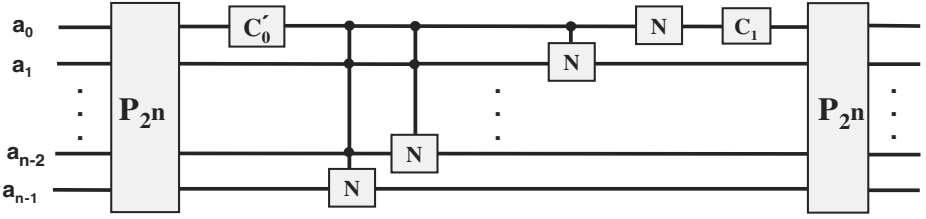


Fig. 10. A circuit for implementation of $D_{2^n}^{(4)}$ by using recursive factorization of Q_{2^n} .

the conditional operators $(\Pi_{2^i} \oplus I_{2^n-2^i})$. However, this leads to a complexity of $O(i^2)$ for implementation of operators $(P_{2^i} \oplus I_{2^n-2^i})$ and hence the operators $(D_{2^i}^{(4)} \oplus I_{2^n-2^i})$. Therefore, while (60) is optimal for implementation of $D_{2^i}^{(4)}$ and the packet algorithm, it is not efficient for implementation of the pyramid algorithm.

It should be emphasized that this recursive factorization of Q_{2^n} , originated by the similarity transformation in (48) and given by (56) and (58), was not previously known in classical computing. Note that, the permutation matrices Π_{2^n} and, particularly, P_{2^n} are much harder (in terms of data movement pattern) for a classical implementation than Q_{2^n} . In this sense, such a factorization of Q_{2^n} is rather counterintuitive from a classical computing point of view since it involves the use of permutation matrices Π_{2^n} and P_{2^n} and thus it is highly inefficient for a classical implementation.

5 Discussion and Conclusion

In this paper, we developed fast algorithms and efficient circuits for quantum wavelet transforms. Assuming an efficient quantum circuit for a given wavelet kernel and starting with a high level description of the packet and pyramid algorithms, we analyzed the feasibility and efficiency of the implementation of the packet and pyramid algorithms by using the given wavelet kernel. We also developed efficient and complete gate-level circuits for two representative wavelet kernels, the Haar and Daubechies $D^{(4)}$ kernels. We gave the first complete time and space complexity analysis of the quantum Haar wavelet transform. We also described three complete circuits for Daubechies $D^{(4)}$ wavelet kernel. In particular, we showed that Daubechies $D^{(4)}$ kernel can be implemented by using the circuit for QFT. Given the problem of decoherence, exploitation of parallelism in quantum computation is a key issue in practical implementation of a given computation. To this end, we are currently analyzing the algorithms of this paper in terms of their parallel efficiency and developing more efficient parallel quantum wavelet algorithms.

As shown in this paper, permutation matrices play a pivotal role in the development of quantum wavelet transforms. In fact, not only they arise explicitly in

the packet and pyramid algorithms but also they play a key role in factorization of wavelet kernels. For classical computing, the implementation of permutation matrices is trivial. However, for quantum computing, it represents a challenging task and demands new, unconventional, and even counterintuitive (from a classical computing view point) techniques. For example, note that most of the factorizations developed in paper for permutation matrices Π_{2^n} , P_{2^n} , and Q_{2^n} were not previously known in classical computing and, in fact, they are not at all efficient for a classical implementation. Also, implementation of the permutation matrices reveals some of the surprises of quantum computing in contrast to classical computing. In the sense that, certain operations that are hard to implement in classical computing are easier to implement in quantum computing and vice versa. As a concrete example, note that while the classical implementation of permutation matrices Π_{2^n} and (particularly) P_{2^n} is much harder (in terms of data movement pattern) than the permutation matrix Q_{2^n} , their quantum implementation is much easier and more straightforward than Q_{2^n} .

In this paper, we focussed on the set of permutation matrices arising in the development of quantum wavelet transforms and analyzed three techniques for their quantum implementation. However, it is clear that the permutation matrices will also play a major role in deriving compact and efficient factorizations, i.e., with polynomial time and space complexity, for other unitary operators by exposing and exploiting their specific structure. Therefore, we believe strongly that a more systematic study of permutation matrices is needed in order to develop further insight into efficient techniques for their implementation in quantum circuits. Such a study might eventually lead to the discovery of new and more efficient approaches for the implementation of unitary transformations and therefore quantum computation.

Acknowledgement

The research described in this paper was performed at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with National Aeronautics and Space Administration (NASA). This work was supported by the NASA/JPL Center for Integrated Space Microsystems (CISM), NASA/JPL Advanced Concepts Office, and NASA/JPL Autonomy and Information Technology Management Program.

Appendix: Commutation of the Operators $I_{2^i} \otimes \Pi_{2^{n-i}}$ with $N \otimes I_{2^{n-j}} \oplus I_{2^{n-2n-j+1}}$

We first prove that every operator of the form $I_{2^i} \otimes \Pi_{2^{n-i}}$, for $i = n - 2$ to 1, commutes with all the operators of the form $N \otimes I_{2^{n-j}} \oplus I_{2^{n-2n-j+1}}$, for $j = i$ to 2, by simply showing that

$$(I_{2^i} \otimes \Pi_{2^{n-i}})(N \otimes I_{2^{n-j}} \oplus I_{2^{n-2(n-j+1)}}) = (N \otimes I_{2^{n-j}} \oplus I_{2^{n-2n-j+1}})(I_{2^i} \otimes \Pi_{2^{n-i}}) \quad (61)$$

The matrix $I_{2^i} \otimes \Pi_{2^{n-i}}$ is a block diagonal matrix and therefore can be written as

$$I_{2^i} \otimes \Pi_{2^{n-i}} = I_2 \otimes \Pi_{2^{n-j}} \oplus I_{2^j-2} \otimes \Pi_{2^{n-j}} \quad (62)$$

It can be then shown that

$$(I_2 \otimes \Pi_{2^{n-j}} \oplus I_{2^j-2} \otimes \Pi_{2^{n-j}})(N \otimes I_{2^{n-j}} \oplus I_{2^{n-2^{n-j}+1}}) = N \otimes \Pi_{2^{n-j}} \oplus I_{2^j-2} \otimes \Pi_{2^{n-j}} \quad (63)$$

and

$$(N \otimes I_{2^{n-j}} \oplus I_{2^{n-2^{n-j}+1}})(I_2 \otimes \Pi_{2^{n-j}} \oplus I_{2^j-2} \otimes \Pi_{2^{n-j}}) = N \otimes \Pi_{2^{n-j}} \oplus I_{2^j-2} \otimes \Pi_{2^{n-j}} \quad (64)$$

It now remains to show that every operator of the form $I_{2^i} \otimes \Pi_{2^{n-i}}$ commutes with the operator $N \otimes I_{2^{n-1}}$. This is simply proved by first using the fact that

$$I_{2^i} \otimes \Pi_{2^{n-i}} = I_2 \otimes (I_{2^{i-1}} \otimes \Pi_{2^{n-i}}) \quad (65)$$

and then showing that

$$\begin{aligned} (I_2 \otimes (I_{2^{i-1}} \otimes \Pi_{2^{n-i}}))(N \otimes I_{2^{n-1}}) &= (N \otimes I_{2^{n-1}})(I_2 \otimes (I_{2^{i-1}} \otimes \Pi_{2^{n-i}})) \\ &= N \otimes I_{2^{i-1}} \otimes \Pi_{2^{n-i}} \end{aligned} \quad (66)$$

References

1. I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, S. Lloyd, "Experimental realization of a quantum algorithm", Nature, 393, p.143, 1998.
2. J. A. Jones, M. Mosca, R. H. Hansen, "Implementation of a Quantum Search Algorithm on a Nuclear Magnetic Resonance Quantum Computer", Nature, 393, p.344, 1998.
3. I. Chuang and Y. Yamamoto, "A Simple Quantum Computer", Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9505011>, 1995.
4. D. Deutsch and R. Jozsa, "Rapid solution of problems by quantum computation", Proc. Royal Society London, Series A, Vol. 439, p. 553, 1992.
5. P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring", Proc. 35th Annual Symposium on Foundations of Computer Science, p. 124, 1994.
6. L. K. Grover, "A Fast Quantum Mechanical Algorithm for Database Search" Proc. 28th Annual ACM Symposium on the Theory of Computing, Philadelphia, p. 212, 1996.
7. Brassard, P. Hoyer, A. Tapp, "Quantum Counting", Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9805082>, 1998.
8. N.J. Cerf, L. K. Grover and C. P. Williams, "Nested quantum search and NP-complete problems", Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9806078>, 1998.
9. W. van Dam, P. Hoyer, A. Tapp, "Multiparty Quantum Communication Complexity", Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9710054>, 1997.
10. C. Zalka, "Grover's quantum searching algorithm is optimal", Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9711070>, 1998.
11. D. Aharonov, A. Kitaev, N. Nisan, "Quantum circuits with mixed states," Proc. 13th Annual ACM Symposium on Theory of Computation, p. 20, 1997.

12. R. Jozsa, "Quantum algorithms and the Fourier transform," Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9707033>, 1997.
13. M. Reck, A. Zeilinger, H.J. Bernstein, and P. Bertani, "Experimental realization of any discrete unitary operator", *Physical Review Letters*, 73, p. 58, 1994.
14. E. Knill, "Approximation by quantum circuits," Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9508006>, 1995.
15. A. Barenco, A. Ekert, K-A Suominen, and P. Torma, "Approximate quantum Fourier transform and decoherence," *Physical Review A*, 54, p. 139, 1996.
16. C. Van Loan, *Computational Frameworks for the Fast Fourier Transform*. SIAM Publications, Philadelphia, 1992.
17. B.J. Fino and R. Alghazi, "A unified treatment of discrete unitary transforms," *SIAM J. Comput.*, 6(4), p. 700, 1977.
18. V. Vedral, A. Barenco, A. Ekert, "Quantum networks for elementary arithmetic operations," *Physical Review A*, 54, p. 147, 1996.
19. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm Pure Appl. Math.*, 41, p. 909, 1988.
20. P. Hoyer, "Efficient quantum Transforms," Los Alamos preprint archive, <http://xxx.lanl.gov/archive/quant-ph/9702028>, Feb. 1997.
21. D. Beckman, A.N. Chari, S. Devabhatuni, and J. Preskill, "Efficient networks for quantum factoring," *Physical Review A*, 54, p. 1034, 1996.
22. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. 2nd Edition, Cambridge Univ. Press, 1992.

Quantum Computation of Fluid Dynamics

Jeffrey Yepez*

Air Force Research Laboratory, Hanscom Field, MA 01731, USA

Abstract. Presented is a quantum lattice gas for Navier-Stokes fluid dynamics simulation. The quantum lattice-gas transport equation at the microscopic scale is presented as a generalization of the classical lattice-gas transport equation. A special type of quantum computer network is proposed that is suitable for implementing the quantum lattice gas. The quantum computer network undergoes a partial collapse of the wavefunction at every time step of the dynamical evolution. Each quantum computer in the network comprises only a few qubits, which are entangled for only a short time period. A Chapman-Enskog type analysis of the quantum computer network indicates that the total system of qubits behaves exactly like a viscous lattice-gas fluid at the macroscopic scale. Because of the quantum mechanical nature of the scattering process, superposition of outgoing collisional possibilities occurs. The quantum lattice gas obeys detail balance in its collisions and is therefore an unconditionally stable algorithm for fluid dynamics simulation.

1 Introduction

Prior to the advent of digital computing in the late 1940's, analog computers held much promise. An electrical circuit can be constructed to simulate, say, an underdamped oscillator governed by a second order differential equation. For example, an electrical circuit can drive a trace on an oscilloscope mimicking the vertical motion response of a fast moving automobile with poor shock absorbers after passing over a speed bump—one continuous physical system configured to behave just like another continuous physical system. Today, after five decades of digital computing, history may repeat itself in the sense that it may again be worthwhile to build “analog” computers—for example, a quantum mechanical spin system configured to behave just like a Navier-Stokes fluid.

The purpose of this paper is to show how to do this. We show how to arrange a network of small quantum computers so that, taken as a system, the qubits within the network mimic the behavior of a system of massive quantum particles moving and colliding on a discrete spacetime lattice. This discrete quantum particle system is termed a *quantum lattice gas* and the associated quantum computer network is called a *lattice-gas quantum computer*.

Over a decade ago, classical lattice gases were found that behave like a viscous Navier-Stokes fluid at the macroscopic scale [12]. In this paper we show that

* This work is supported under Task No. 2304CP of the Air Force Office of Scientific Research, Mathematical and Computational Sciences Directorate.

a quantum lattice gas does too. The prediction of the quantum lattice gas' macroscopic equations of motion is achieved by a generalized Chapman-Enskog analysis. A property of the quantum lattice gas is that continuous macroscopic fields for the mass density and momentum density are obtained without the need for either ensemble averaging or coarse-grain spacetime averaging, both of which are computationally expensive in a classical lattice-gas simulation. This computationally useful property of the quantum lattice gas arises because it models the discrete particle system directly at the mesoscopic scale, avoiding noisy fluctuations while retaining detailed balance in the local particle collisions [3]. Detailed balance is satisfied because of the unitary action of the collision operator as it causes quantum mechanical superpositions of outgoing particle configurations at each site of the spatial lattice. Consequently, the quantum lattice gas is unconditionally stable as a numerical algorithm.

We calculate the single-particle distribution function analytically for a quantum lattice-gas system at local equilibrium. The analytical prediction is that it has the same form as the single-particle distribution function of a classical lattice-gas system, which also obeys the principle of detailed balance. We varyify this prediction through numerical simulation of a two-dimensional quantum lattice gas, which is a straightforward generalization of the classical FHP lattice gas [2]. For comparison purposes, we also include results from a classical FHP simulation. In the low Mach number incompressible fluid regime, there is excellent agreement between the theoretical prediction and the numerical data for the single-particle occupation probability.

2 Review

There are new possibilities and limitations that arise in computing if we use the principle of quantum mechanical superposition of states [4,5,6,7,8,9]. In quantum computing a two-level quantum bit represents the smallest unit of information which may be in a superposition of the discrete states $|0\rangle$ and $|1\rangle$ [1]. An example of the physical embodiment of a qubit is the z-component of a nuclear spin in an atom in a uniform external magnetic field [2].

An open issue for quantum computing is whether or not entangled qubit states (of many qubits, much more than two) within the quantum computer's Hilbert manifold can be isolated from the surrounding environment to a sufficient degree for delicate quantum algorithmic steps to be completed. Using quantum mechanical superposition among qubit states to speedup a computation by simultaneously encoding many possibilities, an approach termed *quantum*

¹ A qubit, $|q\rangle = \alpha|0\rangle + \beta|1\rangle$, has an amplitude, α , of it being in the *zero state*, $|0\rangle$, and another amplitude, β , of finding it in the *one state*, $|1\rangle$. The probabilities add to unity: $\langle 0|0\rangle + \langle 1|1\rangle = \langle q|q\rangle$ so the complex coefficients are constrained by $|\alpha|^2 + |\beta|^2 = 1$.

² Cory *et al.* have employed the quantum number m_z of a nuclear spin of an atom in a molecule of a liquid placed in a strong external magnetic field to encode a single qubit and they used nuclear magnetic resonance to control its state and interaction with qubits in neighboring atoms within the same molecule [10].

parallelism, is generally considered the primary virtue of quantum computation³. Yet uncontrolled coupling with the surrounding environment causes decoherence of the qubit states and the virtue is lost—quantum parallelism levies a high price for coherence of the quantum computer’s wavefunction. This has spurred the development of scalable quantum error correction techniques, considered crucially important for the enterprise to continue [11][8][12][13]. Because of the difficulties of quantum coherence, the first quantum computer comprised only two qubits.

An historical starting point that led to quantum computing was reversible computing [14]. Since microscopic physics is reversible⁴, it is believed that quantum mechanical algorithms must be too.⁵ Reversible algorithms for simulating physics on a quantum device can serve as a guide for constructing the device. The common assumption is the quantum mechanical device itself undergoes unitary (and therefore reversible) evolution as it transitions through its “computation”⁶.

For any reversible computation, one can describe the algorithm by specifying a unitary evolution operator, formally written as $e^{i\hat{H}\tau/\hbar}$, acting on the system wavefunction, $|\Psi\rangle$, which constitutes the state of the quantum computer’s “memory”. With N qubits, the quantum state $|\Psi\rangle$ resides in an exponentially large Hilbert space with 2^N dimensions. A new quantum state, $|\Psi'\rangle$, is generated by application of a unitary matrix of size $2^N \times 2^N$ as follows

$$|\Psi'\rangle = e^{i\hat{H}\tau/\hbar}|\Psi\rangle. \quad (1)$$

By repeated application of $e^{i\hat{H}\tau/\hbar}$ an ordered sequence of states is generated and each one is given a unique time label. If the first state is labeled by t then the next one is labeled by $t + \tau$, and the next by $t + 2\tau$, and so forth. With this understanding we write (II) as

$$|\Psi(t + \tau)\rangle = e^{i\hat{H}\tau/\hbar}|\Psi(t)\rangle. \quad (2)$$

In this way the *computational time* advances incrementally in unit steps of duration τ . Of course the state of the quantum computer exists at all intermediate times, say at $t + \frac{\tau}{2}$, but for our purposes we need only use the state at intervals of the time step τ . The quantum computer’s evolution is invertible by application of the adjoint of the evolution operator

$$|\Psi(t - \tau)\rangle = e^{-i\hat{H}\tau/\hbar}|\Psi(t)\rangle. \quad (3)$$

³ In lattice-gas quantum computation, quantum parallelism is used to allow for simultaneous multiple collision possibilities at each site of the lattice. This allows for a reduction in the viscosity of the fluid which improves the computational efficiency. Yet the computational efficiency is also due to the continuous phase of the qubit, $|q\rangle = \cos\theta|1\rangle + \sin\theta|0\rangle$, which we use to represent the probability of finding a particle, $f_a = \cos^2\theta$.

⁴ Provided photons do not escape to infinity.

⁵ In this paper we use irreversibility, for practical purposes, in part of the quantum mechanical algorithm.

⁶ By restricting oneself to reversible algorithms, in principle heat production may be avoided altogether [15][16].

This computational picture is consistent with the Schrödinger equation of quantum mechanics. For any reversible algorithm chosen, the task is to map the computational evolution operator of the algorithm onto the dynamical evolution of interacting qubits of the physical device.

3 Quantum Lattice Gas

Lattice-gas quantum computation uses the superposition of multiple qubit states within a small spatial region of size ℓ only for a short amount of time on the order of the duration of a single time step, τ . A lattice-gas quantum computer has qubits arranged in a lattice based array with a small group of qubits at each site of the lattice. Each site of the lattice can be thought of as a small quantum computer and all the quantum computers are connected in a lattice network. The quantum lattice gas' evolution can be formally expressed as a special case of (2) as follows

$$|\Psi(\mathbf{x}_1, \dots, \mathbf{x}_V; t + \tau)\rangle = \hat{S}\hat{C} |\Psi(\mathbf{x}_1, \dots, \mathbf{x}_V; t)\rangle, \quad (4)$$

where \hat{S} is a unitary *streaming* matrix and \hat{C} is a unitary *collision* matrix and where we have explicitly labeled the wavefunction's dependence on all the coordinates of the lattice. The operator \hat{C} causes mixing of the outgoing collision configuration at each site of the lattice, locally entangling the qubit states within a lattice cell of radius size, ℓ . The operator \hat{S} causes qubits to move from one site to the next, by exchanges between nearest neighboring sites (it is identical to its classical counterpart). Each qubit moves with unit speed, $c = \frac{\ell}{\tau}$, along one of the lattice directions, \hat{e}_a . Hence, in a completely coherent quantum computation, the application of \hat{S} causes global entanglement of the all the qubit states⁷. It remains an intractable problem to theoretically analyze the dynamics of a quantum computer with many qubits because of the exponentially large size of the Hilbert space in which the entanglement occurs. And to make matters worse, even if a quantum computer was constructed with a large number of qubits, its wavefunction would decohere by uncontrolled entanglement with the external world and we know of no way to mitigate against this. So constructing a large coherent quantum computer is difficult, if not all together impossible, and predicting its behavior by analytical means is intractable.

So what can be done about this? What I would like to consider is a simplification that will sidestep these obstacles and give us two important advantages: (1) a simple way to use a quantum computer with a large number of qubits; and (2) a way to analysis of its behavior. In lattice-gas quantum computation complete coherence of the wavefunction is not needed for the algorithm to work. In fact, we assume entanglement of qubit states is only among small clusters of qubits in a localized nearby neighborhood, so independent quantum operations

⁷ Mathematically speaking, this is because both \hat{S} and \hat{C} cannot be simultaneously diagonalized

are done in a classically parallel fashion on all sites simultaneously. This is the collision step.

In a deterministic classical lattice gas, the collision operator is a permutation matrix with components being either zero or one. In a probabilistic classical lattice gas, the collision operator is a transition matrix with real valued components. In contrast, in a quantum lattice gas, the collision operator can be a unitary matrix with complex components. The collision process is in general irreversible because a projection of the quantum computer's wavefunction into a tensor product state over the qubits is periodically made causing the wavefunction to partially collapse. Hence, application of \hat{S} does not cause any global entanglement.

The quantum lattice gas presented here should not be confused with previous quantum lattice gas models by Succi [17], Boghosian [18], or Yepez [19] for simulating quantum mechanical systems. Despite some similarities, the type of quantum lattice gas treated in this paper is a direct generalization of a classical lattice gas with quantum bits replacing classical bits. In fact, if orthogonal permutation matrices with 0 and ± 1 components are used for the collision process and in the limit of complete collapse of the lattice-gas quantum computer's wavefunction, the quantum lattice gas exactly reduces to a classical lattice gas. This particular feature distinguishes the quantum lattice gas for fluid simulation from the quantum lattice gases for quantum mechanical simulation.

Table 1. Model Constants

Constants	Names
ℓ	length unit
τ	time unit
m	mass unit
c	velocity unit ($\frac{\ell}{\tau}$)
D	spatial dimension
B	lattice coordination number
e_{ai}	unit lattice vectors
a	directional index (1,2,..., B)
i, j, k, l	spatial indices

4 Preliminaries

Consider a lattice-gas quantum computer with the following properties:

- V is the number of lattice sites
- B is the number of qubits per site (and the number of nearest neighbors)
- $N = VB$ is the total number of qubits
- 2^N is the size of the full Hilbert space

Table 2. Wavefunction Symbols

Symbol	Size of Manifold	Description
Ψ	2^N	Total system wavefunction
ψ	2^B	On-site ket
ω	B	Partially collapsed on-site amplitudes
q	2	Qubit ket

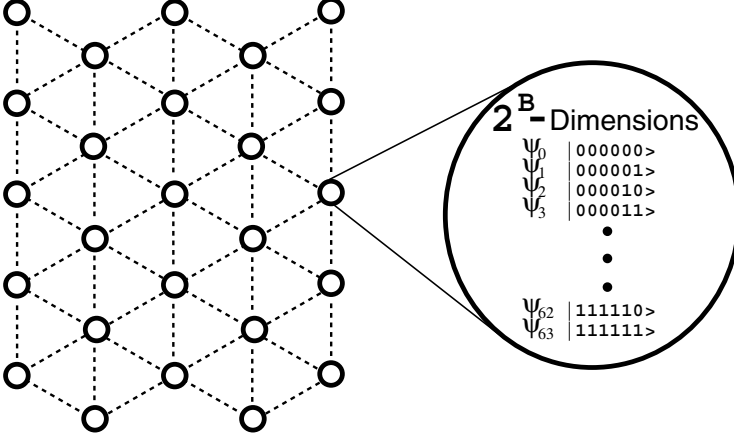


Fig. 1. An array of small quantum computers (the quantum computers are depicted as circles) arranged in a 2-dimensional triangular lattice ($B=6$). The large circle on the right is an expanded view of a single quantum computer which is one site of the lattice. It depicts the on-site submanifold, \mathcal{H} . Each quantum computer at a lattice node has 6 qubits so the on-site ket $|\psi\rangle$ resides in a 64-dimensional Hilbert space. Each node is coupled to its 6 nearest neighboring quantum computers by a mechanism allowing for the exchange of a single qubit.

- 2^B is the size of the on-site submanifold, denoted \mathcal{H}
- B is the size of the reduced on-site submanifold, denoted \mathcal{B}

We will use the following convention for indices:

- Small roman letters (a, b, c) for the \mathcal{B} -space dimensions, $a \in \{1, \dots, B\}$
- Greek letters (α, β, γ) for the \mathcal{H} -space dimensions, $\alpha \in \{0, \dots, 2^B - 1\}$
- Middle roman letters (i, j, k) for the spatial dimensions, $i \in \{1, \dots, D\}$

The full Hilbert space of size 2^{BV} is partitioned into V independent quantum manifolds of dimension 2^B , as depicted in Fig. 1. Quantum superposition of states occurs only within each 2^B -dimensional subspace, denoted \mathcal{H} . A general *on-site ket* defined over the basis states of \mathcal{H} is the following

$$|\psi(\mathbf{x}, t)\rangle = \sum_{\alpha=0}^{2^B-1} \psi_{\alpha}(\mathbf{x}, t) |\alpha\rangle = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{2^B-1} \end{pmatrix}. \quad (5)$$

The ket $|\psi\rangle$ is specified by 2^B complex amplitudes, denoted $\psi_0, \dots, \psi_{2^B-1}$. The quantum computer's total wavefunction is formed as a tensor product over all the \mathcal{H} -manifolds

$$|\Psi(\mathbf{x}_1, \dots, \mathbf{x}_V; t)\rangle = \bigotimes_{x=1}^V |\psi(\mathbf{x}, t)\rangle. \quad (6)$$

The collision operator, \hat{C} , is blocked over all the \mathcal{H} -manifolds. That is, the collision matrix is block diagonal with V blocks each of size $2^B \times 2^B$, and therefore can be written as a tensor product

$$\hat{C} = \bigotimes_{x=1}^V \hat{U}. \quad (7)$$

The *on-site collision matrix*, \hat{U} , is unitary and acts on the on-site ket

$$|\psi'(\mathbf{x}, t)\rangle = \hat{U} |\psi(\mathbf{x}, t)\rangle. \quad (8)$$

The prime on the L.H.S. of (8) indicates that the ket is an *outgoing* collisional state.

5 Unitary Collision Matrix

Let \hat{Q}_{α} be matrices representing the conserved quantities in the single speed quantum lattice gas, $\hat{Q}_{\alpha} = (\hat{Q}_{\circ}, \hat{Q}_i)$ where i is an index over the independent spatial coordinates. A fundamental property of a quantum lattice gas is that the mass density and the momentum density can be written as follows

$$\rho = \langle \psi | \hat{Q}_{\circ} | \psi \rangle \quad (9)$$

$$\rho v_i = \langle \psi | \hat{Q}_i | \psi \rangle, \quad (10)$$

that is, where the component of \hat{Q}_{\circ} are

$$(\hat{Q}_{\circ})_{\mu\nu} \equiv m \delta_{\mu\nu} \sum_{a=1}^B b_{\mu a}, \quad (11)$$

and where the component of \hat{Q}_i are

$$(\hat{Q}_i)_{\mu\nu} \equiv m c \delta_{\mu\nu} \sum_{a=1}^B b_{\mu a} e_{ai}. \quad (12)$$

The $b_{\mu a}$ denotes the a^{th} -bit of the μ^{th} ket. The \hat{e}_a here denote the unit lattice vectors where $a = 1, \dots, B$. The components of the $2^B \times 2^B$ qubit number operator are defined by

$$(\hat{n}_a)_{\mu\nu} \equiv b_{\mu a} \delta_{\mu\nu}. \quad (13)$$

In terms of (13), the operators for mass and momentum are

$$\hat{Q}_o = m \sum_{a=1}^B \hat{n}_a \quad (14)$$

and

$$\hat{Q}_i = mc \sum_{a=1}^B e_{ai} \hat{n}_a. \quad (15)$$

In terms of (13) the invariant quantities are simply expressed as the following matrix elements

$$\rho = \sum_{a=1}^B m \langle \psi | \hat{n}_a | \psi \rangle \quad (16)$$

$$\rho v_i = \sum_{a=1}^B m c e_{ai} \langle \psi | \hat{n}_a | \psi \rangle. \quad (17)$$

The matrix elements (9) and (10) must remain constant after each time step iteration

$$\langle \psi(t + \tau) | \hat{Q}_\alpha | \psi(t + \tau) \rangle = \langle \psi(t) | \hat{Q}_\alpha | \psi(t) \rangle. \quad (18)$$

Since $|\psi(t + \tau)\rangle = \hat{U} |\psi(t)\rangle$, this implies that

$$\hat{U}^\dagger \hat{Q}_\alpha \hat{U} = \hat{Q}_\alpha, \quad (19)$$

which is just the commutator

$$[\hat{U}, \hat{Q}_\alpha] = 0. \quad (20)$$

The matrices \hat{Q}_α must commute with \hat{U} .

Let \hat{g} denote the generator of \hat{U}

$$\hat{U} = e^{i\varepsilon \hat{g}}, \quad (21)$$

where ε is an ‘‘Euler angle’’. Consider a ‘‘rotation’’ through an infinitesimal angle ε so that \hat{U} can be Taylor expanded to first order as

$$\hat{U} = \mathbf{1} + i\varepsilon \hat{g}. \quad (22)$$

The unitary condition, $\hat{U}^\dagger \hat{U} = \mathbf{1}$, implies that the generator is hermitian

$$\hat{g} - \hat{g}^\dagger = 0 + \mathcal{O}(\varepsilon^2). \quad (23)$$

From (19), we see that mass and momentum conservation is ensured provided

$$\hat{Q}_\alpha \hat{g} - \hat{g}^\dagger \hat{Q}_\alpha = 0 + \mathcal{O}(\varepsilon^2). \quad (24)$$

The solution of the set of linear equations (23) and (24) give the Lie algebra for the unitary group. Therefore, the mass density (9) and the momentum density (10) are conserved when each equivalence class block of the collision operator is an element of the unitary group $U(n)$ where n is the size of the equivalence class of the incoming local configuration. This is an important feature of a quantum lattice gas. Since any member of the unitary group can be used, the quantum lattice gas is algorithmically robust.

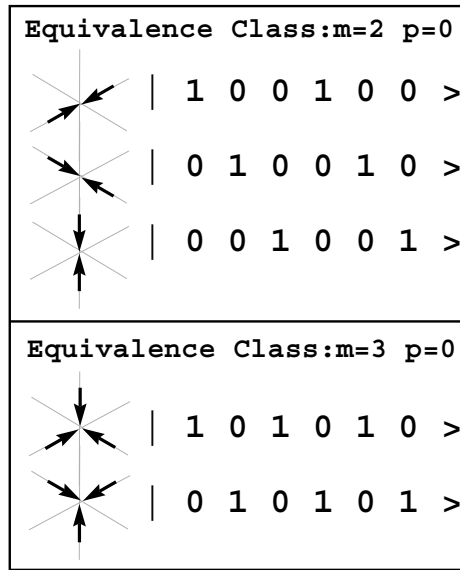


Fig. 2. The equivalence classes for the quantum FHP lattice gas.

An *equivalence class* is defined as a set of basis states that correspond to particle configurations with the same mass and momentum. The unitary collision operator, \hat{U} , acting on the 2^B dimensional \mathcal{H} -manifold itself is block diagonal over the equivalence classes. For example, there are two equivalence classes for the FHP lattice gas [2], see Fig 2. The first equivalence class is comprised of the following two-body kets

$$\begin{aligned} |9\rangle &= |001001\rangle \\ |18\rangle &= |010010\rangle \\ |36\rangle &= |100100\rangle \end{aligned}$$

with mass, $m = 2$, and zero momentum, $\mathbf{p} = 0$. A general ket in this mass-momentum sector of the on-site manifold is a linear combination of these

$$\alpha|100100\rangle + \beta|010010\rangle + \gamma|001001\rangle, \quad (25)$$

where α , β , and γ are complex numbers. The second equivalence class is comprised of the following three-body kets

$$\begin{aligned} |21\rangle &= |010101\rangle \\ |42\rangle &= |101010\rangle \end{aligned}$$

with mass, $m = 3$, and zero momentum, $\mathbf{p} = 0$. A general ket in this mass-momentum sector is a linear combination of these

$$\mu|101010\rangle + \nu|010101\rangle. \quad (26)$$

So \hat{U} for a two-dimensional quantum lattice gas on a triangular lattice has two blocks, a $U(3)$ block for mixing the 2-body configurations and a $U(2)$ block for mixing the 3-body configurations.

For the triangular quantum lattice gas, we have

$$\begin{pmatrix} \psi'_{21} \\ \psi'_{42} \end{pmatrix} = e^{i\theta} \begin{pmatrix} e^{i\zeta} \cos \eta & e^{i\zeta} \sin \eta \\ -e^{-i\zeta} \sin \eta & e^{-i\zeta} \cos \eta \end{pmatrix} \begin{pmatrix} \psi_{21} \\ \psi_{42} \end{pmatrix}, \quad (27)$$

where zero momentum three-body configurations are mixed by a unitary matrix, $U(2) = U(1) \otimes SU(2)$, which in general has four free parameters. The zero momentum two-body configurations are mixed by a unitary matrix, $U(3) = U(1) \otimes SU(3)$, which in general has nine free parameters⁸

$$\begin{pmatrix} \psi'_9 \\ \psi'_{18} \\ \psi'_{36} \end{pmatrix} = e^{i\theta} SU(3) \begin{pmatrix} \psi_9 \\ \psi_{18} \\ \psi_{36} \end{pmatrix}. \quad (28)$$

6 Partial Collapse of Post-Collision Ket $|\psi'\rangle$

To avoid causing any global entanglement as induced by streaming, we project the post-collision ket $|\psi'\rangle$ which resides in the 2^B -dimensional \mathcal{H} manifold onto a smaller B -dimensional submanifold, \mathcal{B} , using a projection operator, denoted \hat{F} , as follows

$$|\omega'\rangle = \hat{F} |\psi'\rangle = \begin{pmatrix} \omega'_1 \\ \omega'_2 \\ \vdots \\ \omega'_B \end{pmatrix}. \quad (29)$$

The operator \hat{F} causes a partial collapse of the locally entangled on-site state $|\psi\rangle$ resulting in a nonentangled state $|\omega\rangle$ residing in a smaller manifold (a mapping

⁸ We do not write out the $SU(3)$ matrix in component form because it is too complicated.

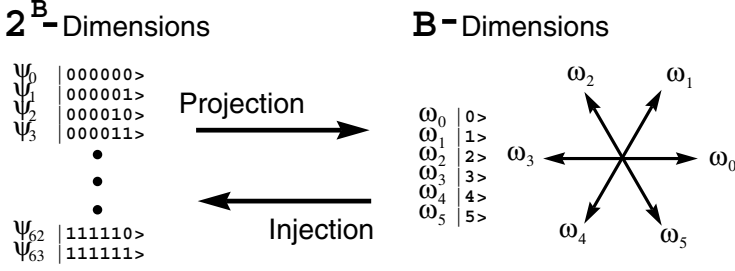


Fig. 3. Two-dimensional quantum lattice gas on a triangular lattice. The lattice coordination number is $B = 6$. There are $2^6 = 64$ amplitudes in the \mathcal{H} -manifold and 6 amplitudes in the \mathcal{B} -manifold. *Injection* maps the 6 on-site amplitudes ω_a in the \mathcal{B} -manifold into the larger \mathcal{H} -manifold. The inverse process, *projection*, maps the 64 amplitudes ψ_α in the \mathcal{H} -manifold onto 6 amplitudes ω_a in the \mathcal{B} -manifold. The projection is a measurement process that causes a partial collapse of the on-site wavefunction $|\psi(\mathbf{x}, t)\rangle$. The partially collapsed wavefunction is $|\omega(\mathbf{x}, t)\rangle$.

from 64 dimensions down to 6, see Fig. 3. Thus $|\omega\rangle$ in (29) may be termed the *collapsed post-collisional on-site ket*. By construction, the action of \hat{F} fixes the phase, θ_a , of the on-site qubits $|q_a\rangle$ according to the following recipe

$$|q_a\rangle = \cos \theta_a |1\rangle + \sin \theta_a |0\rangle, \quad (30)$$

where $\omega_a = \cos \theta_a$. That is,

$$|q_a\rangle = \omega_a |1\rangle + \sqrt{1 - \omega_a^* \omega_a} |0\rangle. \quad (31)$$

After the collapse of the ket $|\psi\rangle$ the *single-particle occupation probability*, denoted f_a , is a well-defined quantity. It is the probability of finding a particle at coordinate (\mathbf{x}, t) with momentum $m\mathbf{c}\hat{e}_a$

$$f_a(\mathbf{x}, t) = \omega_a^*(\mathbf{x}, t) \omega_a(\mathbf{x}, t). \quad (32)$$

Using the single qubit number operator $\hat{n} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, (32) can be written in terms of the qubit ket as

$$f_a(\mathbf{x}, t) = \langle q_a(\mathbf{x}, t) | \hat{n} | q_a(\mathbf{x}, t) \rangle. \quad (33)$$

Furthermore, in (16) and (17), the matrix element $\langle \psi | \hat{n}_a | \psi \rangle$ also gives the probability of particle occupancy. So for a quantum lattice gas, f_a can also be expressed as the matrix element of the multiple qubit number operator⁹

$$f_a = \langle \psi | \hat{n}_a | \psi \rangle. \quad (34)$$

⁹ In a classical lattice gas the single-particle occupation probability is obtained by ensemble averaging over the number variables, $f_a = \langle n_a \rangle$, where $n_a = 0$ or 1.

Inserting the expression for the outgoing collisional state (8) into the R.H.S. of (29), we have

$$| \omega' \rangle = \hat{F} \hat{U} | \psi \rangle. \quad (35)$$

In §7 we will use a nonlinear function for the projection operator. Our goal is to retain as much quantum information as possible while allowing (35) to reduce to the collision equation of classical lattice gas transport when \hat{U} is a real-valued permutation matrix. This is accomplished by projecting down from the \mathcal{H} -manifold containing 2^B complex amplitudes to the \mathcal{B} -manifold with only B complex amplitudes. Each ω_a (or associated qubit $| q_a \rangle$) is “attached” to one of the lattice directions. The reason for this reduction of the quantum information is that by using only B complex amplitudes, one for each direction, we can straightforwardly write down a quasi-classical streaming equation in analogy to the streaming equation of a classical lattice gas

$$| \omega(\mathbf{x} + \ell \hat{\mathcal{L}}_a, t + \tau) \rangle = | \omega'(\mathbf{x}, t) \rangle, \quad (36)$$

where $\mathcal{L}_{abi} \equiv \hat{e}_{ai} \delta_{ab}$. Inserting (35) into (36), we have

$$| \omega(\mathbf{x} + \ell \hat{\mathcal{L}}_a, t + \tau) \rangle = \hat{F} \hat{U} | \psi(\mathbf{x}, t) \rangle. \quad (37)$$

The only task remaining to complete the analogy to classical lattice gas dynamics is to rewrite the R.H.S. of (37) solely in terms of the ω 's. This can be done by *injecting* the on-site collapsed ket $| \omega \rangle$ residing in the sub-manifold \mathcal{B} up into the larger on-site manifold \mathcal{H} (see Fig 3). This process can be expressed by application of an *injection operator*, \hat{I} , as follows

$$| \psi(\mathbf{x}, t) \rangle = \hat{I} | \omega(\mathbf{x}, t) \rangle. \quad (38)$$

A straightforward way to accomplish the injection is to take the tensor product over the on-site qubits

$$| \psi(\mathbf{x}, t) \rangle = \bigotimes_{a=1}^B | q_a(\mathbf{x}, t) \rangle. \quad (39)$$

This is a nonlinear operation.¹⁰ Let us revisit the example a two-dimensional quantum lattice gas on a triangular lattice ($B = 6$), a generalization of the classical FHP lattice gas [2]. Fig 3 illustrates projection from the 64-dimensional \mathcal{H} -manifold down to the 6-dimensional \mathcal{B} -manifold and illustrates injection from the \mathcal{B} -manifold up to the \mathcal{H} -manifold. The 6 on-site amplitudes ω_a (or the associated 6 on-site qubits $| q_a \rangle$) generated by the projection can be streamed in a classical fashion. After streaming to their new sites, each quantum computer has a new incoming configuration of the ω_a amplitudes. Before this configuration

¹⁰ A non-square linear matrix could also be used, but it is difficult to find an appropriate matrix even though it can be shown that one exists.

can be collided, they must be injected up to the larger 64-dimensional manifold where the collision process is well defined.

Inserting (38) into (35) gives

$$|\omega'(\mathbf{x}, t)\rangle = \hat{F}\hat{U}\hat{I} |\omega(\mathbf{x}, t)\rangle. \quad (40)$$

Using the fact that the projection of the injection is the identity operation: $|\omega\rangle \equiv \hat{F}\hat{I} |\omega\rangle$, we write (40) in a form analogous to the classical lattice gas collision equation

$$|\omega'(\mathbf{x}, t)\rangle = |\omega(\mathbf{x}, t)\rangle + \left[\hat{F}\hat{U}\hat{I} |\omega(\mathbf{x}, t)\rangle - \hat{F}\hat{I} |\omega(\mathbf{x}, t)\rangle \right]. \quad (41)$$

Finally, we arrive at the quantum lattice gas microscopic transport equation by inserting (41) into (36)

$$|\omega(\mathbf{x} + \ell\hat{\mathcal{L}}_a, t + \tau)\rangle = |\omega(\mathbf{x}, t)\rangle + |\Omega(\mathbf{x}, t)\rangle, \quad (42)$$

where the quantum lattice gas collision operator is defined as

$$|\Omega(\mathbf{x}, t)\rangle \equiv \hat{F}\hat{U}\hat{I} |\omega(\mathbf{x}, t)\rangle - \hat{F}\hat{I} |\omega(\mathbf{x}, t)\rangle. \quad (43)$$

In component form, (42) is

$$\omega_a(\mathbf{x} + \ell\hat{e}_a, t + \tau) = \omega_a(\mathbf{x}, t) + \Omega_a(\omega_*(\mathbf{x}, t)). \quad (44)$$

Equation (44) is identical in form to the classical lattice gas transport equation where the occupation variable, $n_a = 0$ or 1 , is replaced by a complex amplitude, $0 \leq |\omega_a| \leq 1$, that continuously encodes the square root of the probability for particle occupancy. Hence (44) is a much more useful expression of the quantum lattice gas dynamics than (4) is.

7 The Projection Operator

We can write an analytical expression for the projection operator where the amplitudes ω_a a nonlinear function of the amplitudes ψ_a

$$\omega_a = \hat{I}_a(\psi) \equiv \sqrt{\sum_{\alpha=0}^{2^B-1} |\psi_\alpha|^2 b_{\alpha a}}. \quad (45)$$

Note that $b_{\alpha a} = 0$ or 1 is the Boolean value of the a^{th} bit of the α^{th} ket in the number representation. Let \hat{m} and \hat{p}_i be the operators for mass and momentum in the \mathcal{B} -space. Then the matrix element for the mass density is

$$\rho = \langle \omega | \hat{m} | \omega \rangle, \quad (46)$$

where $\hat{m}_{ab} = m\delta_{ab}$, and the matrix element for the momentum density is

$$\rho v_i = \langle \omega | \hat{p}_i | \omega \rangle, \quad (47)$$

where $(\hat{p}_i)_{ab} = mce_{ai}\delta_{ab}$. \hat{Q}_o and \hat{Q}_i were defined in §5 to be the mass and momentum operators in ψ -space. Here we have mass and momentum operators in ω -space. The matrix element (9) defines the mass density as $\rho = \langle \psi | \hat{Q}_o | \psi \rangle$, where $(\hat{Q}_o)_{\alpha\beta} = m \sum_{a=1}^B b_{a\alpha} \delta_{\alpha\beta}$, and the matrix element (10) defines the momentum density as $\rho v_i = \langle \psi | \hat{Q}_i | \psi \rangle$, where $(\hat{Q}_i)_{\alpha\beta} = mc \sum_{a=1}^B b_{a\alpha} e_{ai} \delta_{\alpha\beta}$. Equating (46) with (9) and equating (47) with (10) gives us a way to check the projection operator (45).

This is done as follows

$$\langle \omega | \hat{m} | \omega \rangle = m \sum_{a=1}^B |\omega_a|^2 \quad (48)$$

$$= m \sum_{a=1}^B \sum_{\alpha=1}^{2^B} |\psi_\alpha|^2 b_{\alpha a} \quad (49)$$

$$= \sum_{\alpha=1}^{2^B} |\psi_\alpha|^2 \left(m \sum_{a=1}^B b_{\alpha a} \right), \quad (50)$$

where we used the square of (45) on the second line of the derivation. Therefore, we have

$$\langle \omega | \hat{m} | \omega \rangle = \langle \psi | \hat{Q}_o | \psi \rangle, \quad (51)$$

where the mass operator in ψ -space

$$(\hat{Q}_o)_{\alpha\beta} = m \sum_{a=1}^B b_{\alpha a} \delta_{\alpha\beta} \quad (52)$$

is identical to \hat{Q}_o defined in (9). So the projection operator (45) conserves mass. We continue the consistency check by rewriting (47)

$$\langle \omega | \hat{p}_i | \omega \rangle = mc \sum_{a=1}^B |\omega_a|^2 e_{ai} \quad (53)$$

$$= mc \sum_{a=1}^B \sum_{\alpha=1}^{2^B} |\psi_\alpha|^2 b_{\alpha a} e_{ai} \quad (54)$$

$$= \sum_{\alpha=1}^{2^B} |\psi_\alpha|^2 \left(mc \sum_{a=1}^B b_{\alpha a} e_{ai} \right), \quad (55)$$

where again we used the square of (45) on the second line of the derivation. Therefore, we have

$$\langle \omega | \hat{p}_i | \omega \rangle = \langle \psi | \hat{Q}_i | \psi \rangle, \quad (56)$$

where the momentum operator in ψ -space

$$(\hat{Q}_i)_{\alpha\beta} = mc \sum_{a=1}^B b_{\alpha a} e_{ai} \delta_{\alpha\beta}, \quad (57)$$

is identical to \hat{Q}_i defined in (10). So the projection operator (45) conserves momentum as well as mass.

8 Equilibrium Ansatz

The ω_a amplitudes in \mathcal{B} -space can be ordered in powers of ε as follows

$$|\omega\rangle = |\omega^{(0)}\rangle + \varepsilon |\omega^{(1)}\rangle + \mathcal{O}(\varepsilon^2), \quad (58)$$

where $|\omega^{(0)}\rangle$ denotes the *equilibrium ket* and where the ket $|\omega^{(1)}\rangle$ is the first order correction from equilibrium. The condition equilibrium is that $|\omega^{(0)}\rangle$ satisfies the following identity

$$|\omega^{(0)}\rangle = \hat{F}\hat{U}\hat{F} |\omega^{(0)}\rangle. \quad (59)$$

Note that the associated equilibrium ket in \mathcal{H} -space, $|\psi^{(0)}\rangle$, follows from (58) by injection $|\psi\rangle = \hat{I} |\omega\rangle$

$$|\psi\rangle = |\psi^{(0)}\rangle + \varepsilon |\psi^{(1)}\rangle + \mathcal{O}(\varepsilon^2). \quad (60)$$

So we can also write (59) as follows

$$|\psi^{(0)}\rangle = \hat{U} |\psi^{(0)}\rangle. \quad (61)$$

It is clear that $|\psi^{(0)}\rangle$ is an eigenvector of \hat{U} with unity eigenvalue. Using (59), we immediately see that the collision operator (43) vanishes at equilibrium

$$|\Omega(\omega^{(0)})\rangle = \hat{F}\hat{U}\hat{F} |\omega^{(0)}\rangle - \hat{F}\hat{I} |\omega^{(0)}\rangle = 0. \quad (62)$$

Equations (58) and (59) constitute the essential ansatz that will allow us to perform a Chapman-Enskog analysis of the quantum lattice gas. It is possible to analytically solve (59) for $|\omega^{(0)}\rangle$. Knowing the form of $|\omega^{(0)}\rangle$, we can predict the hydrodynamics equations of the quantum lattice gas at the macroscopic scale. In the Chapman-Enskog analysis, we expand the collision operator, $|\Omega\rangle$, about this equilibrium ket $|\omega^{(0)}\rangle$. In so doing, the Jacobian of the collision operator is computed as a first order correction and is evaluated at $|\omega\rangle = |\omega^{(0)}\rangle$. Since the transport coefficients for the mass diffusion, shear viscosity, and bulk viscosity depend on the value of this first order correction, and this in turn depends on the value of $|\omega^{(0)}\rangle$, one must determine the equilibrium amplitudes in order to compute the value of the transport coefficients.

The a single particle occupancy probability $f_a = \omega_a^{*(0)} \omega_a^{(0)} = \langle q_a^{(0)} | \hat{n} | q_a^{(0)} \rangle = \langle \psi | \hat{n}_a | \psi \rangle$ has the functional form

$$f_a = \frac{1}{e^{\alpha\rho + \beta\vec{e}_a \cdot \vec{p} + \gamma E} + 1}, \quad (63)$$

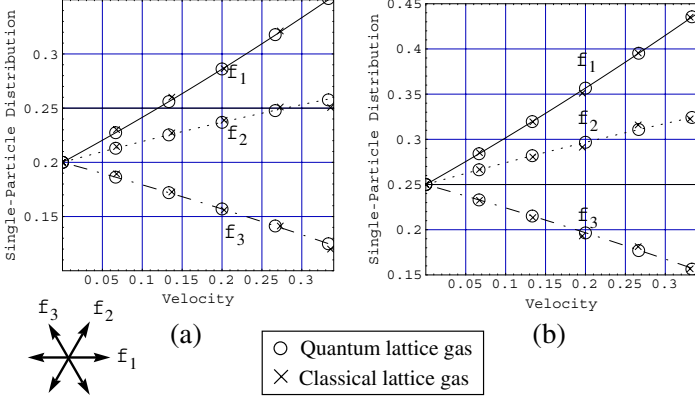


Fig. 4. Theory versus simulation comparison of the velocity dependence of the single-particle distribution function in the non-Galilean parameterization: $f_a = \omega_a^* \omega_a = d + dD\hat{e}_a \cdot \mathbf{v} + gdD(D/2 + 1)\hat{Q}_a : \mathbf{v}\mathbf{v}$. FHP simulation data is overplotted on this predicted mesoscopic distribution function. Plots (a) and (b) are for background densities of $d = .20$ and $d = 0.25$, respectively. A velocity shift is imparted along the x -axis; that is, along the f_1 direction indicated in the figure. Data was collected from a 128×128 classical FHP simulation (crosses) and was coarse-grained averaged over 1600 time steps from time step $t = 400$ to $t = 2000$. Data was also collected from a smaller 32×32 quantum FHP simulation (circles) and were measured at a single time step at $t = 200$.

where the argument of the exponential is a linear combination of the conserved scalar quantities: (1) the mass ρ ; (2) the momentum component $\hat{e}_a \cdot \mathbf{p}$ along the lattice direction \hat{e}_a ; and (3) the energy E at a lattice site. The real valued coefficients α , β , and γ are free parameters that are fixed by the non-Galilean parameterization given in Appendix A. The reason for the form of (63) is the collision matrix \hat{U} is unitary and so the collisions obey detailed balance.

In the quantum limit, where the quantum lattice gas becomes a fully coherent quantum system that undergoes unitary evolution, we expect the equilibrium probability for the particle occupancy to have the form of (63). In the opposite limit, the classical limit, where there is a complete collapse of the ket $|\Psi\rangle$ everywhere, the quantum lattice gas reduces to a classical lattice gas system. In the classical limit too, the particle occupancies are described by (63). Our quantum lattice gas dynamics is somewhere midway between a fully quantum system and a fully classical system. Since both ends of the spectrum are described by (63), it is not altogether unexpected that the middle regime is too. However, this is not obvious. At first glance, it appears that the destructive act of projecting from the \mathcal{H} -manifold onto the \mathcal{B} -manifold might destroy the form of the equilibrium distribution. But it does not (see Fig 4); the distribution (63) describes the particle occupancy of our quantum lattice gas [3]. Taking $\omega_a = \sqrt{\langle q_a^{(0)} | \hat{n} | q_a^{(0)} \rangle}$, the result (see Appendix A) is the following¹¹

¹¹ In (64), we have used the approximation $\sqrt{1+x} \cong 1 + \frac{x}{2}$, which holds for small x .

$$|\omega\rangle = \sqrt{d} \left(|\mathbf{1}\rangle + \frac{D}{2c} \mathbf{v} \cdot |\hat{e}\rangle + \frac{gD(D+2)}{4c^2} \mathbf{v} \mathbf{v} : |Q\rangle \right), \quad (64)$$

where

$$|\mathbf{1}\rangle \equiv \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (65)$$

$$|\hat{e}\rangle \equiv \hat{\mathcal{L}} |\mathbf{1}\rangle, \quad (66)$$

and

$$|Q\rangle \equiv (\hat{\mathcal{L}}\hat{\mathcal{L}} - \frac{\mathbf{1}}{D}) |\mathbf{1}\rangle. \quad (67)$$

Taylor expanding the collision operator gives

$$\Omega_a(\omega) = \Omega_a(\omega^{(0)}) + \varepsilon \frac{\partial \Omega_a(\omega)}{\partial \omega_b} \Big|_{\omega=\omega^{(0)}} \omega_b^{(1)} + \mathcal{O}(\varepsilon^2). \quad (68)$$

Using the equilibrium condition (62), the first term on the R.H.S. vanishes and we are left with

$$\Omega_a(\omega) = \varepsilon \mathcal{J}_{ab} \omega_b^{(1)} + \mathcal{O}(\varepsilon^2), \quad (69)$$

where the Jacobian of the projection operator is defined as

$$\mathcal{J}_{ab} \equiv \frac{\partial \Omega_a(\omega)}{\partial \omega_b} \Big|_{\omega=\omega^{(0)}}. \quad (70)$$

Equation (69) can be written in vector form as

$$|\hat{\Omega}(\omega)\rangle = \varepsilon \mathcal{J} |\omega^{(1)}\rangle + \mathcal{O}(\varepsilon^2). \quad (71)$$

9 Boltzmann Equation for the Quantum Lattice Gas

The quantum lattice gas microscopic transport equation (42) is

$$|\omega(\mathbf{x}\mathbf{1} + \varepsilon \ell \hat{\mathcal{L}}, t + \varepsilon^2 \tau)\rangle = |\omega(\mathbf{x}, t)\rangle + |\Omega(\mathbf{x}, t)\rangle.$$

Near equilibrium, each application of the projection operator causes only small changes in the phase of $|\omega\rangle$ allowing us to Taylor expand the L.H.S. of (42). Taylor expanding (42) to second order in ε gives a Boltzmann equation

$$\varepsilon^2 \partial_t |\omega(\mathbf{x}, t)\rangle + \left(\varepsilon c \hat{\mathcal{L}}_i \partial_i + \frac{\varepsilon^2 \ell^2}{2\tau} \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j \partial_i \partial_j \right) |\omega(\mathbf{x}, t)\rangle + \mathcal{O}(\varepsilon^3) = \frac{1}{\tau} |\Omega(\mathbf{x}, t)\rangle. \quad (72)$$

In analogy with a classical lattice, we impose two constraints on the collision operator regarding the isometries of the lattice. The first constraint is

$$\langle \omega | \hat{m} | \Omega \rangle = 0, \quad (73)$$

and this will be needed when we take the zeroth moment of (72). This constraint enforces mass conservation. The second constraint to enforce momentum conservation is

$$\langle \omega | \hat{p}_i | \Omega \rangle = 0, \quad (74)$$

and this will be needed when we take the first moment of (72). Inserting the ε -expansion of $|\omega_a\rangle$, (58), and the ε -expansion of $|\Omega_a\rangle$, (71), into the quantum lattice gas transport equation, (72), and keeping terms up to second order in ε gives

$$\begin{aligned} \varepsilon^2 \partial_t |\omega^{(0)}\rangle + \frac{\varepsilon}{\tau} \hat{\mathcal{L}}_i \partial_i |\omega^{(0)}\rangle + \frac{\varepsilon^2}{\tau} \hat{\mathcal{L}}_i \partial_i |\omega^{(1)}\rangle + \frac{\varepsilon^2}{2\tau} \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j \partial_i \partial_j |\omega^{(0)}\rangle = \\ \frac{\varepsilon}{\tau} \mathcal{J} |\omega^{(1)}\rangle + \mathcal{O}(\varepsilon^3). \end{aligned} \quad (75)$$

Now equating the order- ε terms, we get

$$\hat{\mathcal{J}} |\omega^{(1)}\rangle = \hat{\mathcal{L}}_i \partial_i |\omega^{(0)}\rangle. \quad (76)$$

Then inverting the kinetic part of the Jacobian matrix gives

$$|\omega^{(1)}\rangle = \ell \hat{\mathcal{J}}^{-1} \left(\omega^{(0)} \right) \hat{\mathcal{L}}_i \partial_i |\omega^{(0)}\rangle. \quad (77)$$

So the first order correction ket is equated to the gradient of the equilibrium ket. Inserting this result back into (58) we have

$$|\omega\rangle = |\omega^{(0)}\rangle + \varepsilon \hat{\mathcal{J}}^{-1} \hat{\mathcal{L}}_i \partial_i |\omega^{(0)}\rangle + \mathcal{O}(\varepsilon^2), \quad (78)$$

Inserting (64) into (78), we have the subsonic ε -expansion of $|\omega\rangle$ good to first order in ε

$$\begin{aligned} |\omega\rangle = \sqrt{d} \left(|1\rangle + \frac{D}{2c} \mathbf{v} \cdot \hat{\mathbf{e}} + \frac{gD(D+2)}{4c^2} \mathbf{v} \mathbf{v} : |Q\rangle + \frac{\varepsilon D}{2c} \hat{\mathcal{J}}^{-1} \hat{\mathcal{L}} \hat{\mathcal{L}} : \nabla \mathbf{v} |1\rangle \right) + \\ \mathcal{O}(\varepsilon^2). \end{aligned} \quad (79)$$

10 Hydrodynamic Equations

ZEROth MOMENT EQUATION:

Now we can write the *zeroth moment* of the Boltzmann equation (72) by left multiplying by $\langle \omega | \hat{m}$ as follows

$$\varepsilon^2 \langle \omega | \hat{m} | \partial_t \omega \rangle + \varepsilon \langle \omega | \hat{m} c \hat{\mathcal{L}}_i | \partial_i \omega \rangle + \frac{\varepsilon^2 \ell^2}{2\tau} \langle \omega | \hat{m} \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \partial_i \partial_j \omega \rangle = \frac{1}{\tau} \langle \omega | \hat{m} | \Omega \rangle. \quad (80)$$

The adjoint of (80) is the following

$$\varepsilon^2 \langle \partial_t \omega | \hat{m} | \omega \rangle + \varepsilon \langle \partial_i \omega | \hat{\mathcal{L}}_i \hat{m} c | \omega \rangle + \frac{\varepsilon^2 \ell^2}{2\tau} \langle \partial_i \partial_j \omega | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j \hat{m} | \omega \rangle = \frac{1}{\tau} \langle \Omega | \hat{m} | \omega \rangle, \quad (81)$$

and the R.H.S. will also vanishes because of constraint (73). Because \hat{m} and \mathcal{L} commute, adding (80) and (81) gives us the following zeroth moment equation

$$\varepsilon^2 \partial_t \langle \omega | \hat{m} | \omega \rangle + \varepsilon \partial_i \langle \omega | \hat{m} c \hat{\mathcal{L}}_i | \omega \rangle + \frac{\varepsilon^2 \ell^2}{2\tau} \left(\langle \omega | \hat{m} \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \partial_i \partial_j \omega \rangle + \langle \partial_i \partial_j \omega | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j \hat{m} | \omega \rangle \right) = 0. \quad (82)$$

FIRST MOMENT EQUATION:

Now we can write the *first moment* of (72) by left multiplying by $\langle \omega | \hat{p}_i$ as follows

$$\varepsilon^2 \langle \omega | \hat{p}_i | \partial_t \omega \rangle + \varepsilon \langle \omega | \hat{p}_i c \hat{\mathcal{L}}_j | \partial_j \omega \rangle + \frac{\varepsilon^2 \ell^2}{2\tau} \langle \omega | \hat{p}_i \hat{\mathcal{L}}_j \hat{\mathcal{L}}_k | \partial_j \partial_k \omega \rangle = 0, \quad (83)$$

where the R.H.S. vanishes because of the second constraint (74) on the collision operator. Because \hat{p} and \mathcal{L} commute, adding (83) to its adjoint equation gives us the following first moment equation

$$\varepsilon^2 \partial_t \langle \omega | \hat{p}_i | \omega \rangle + \varepsilon \partial_j \langle \omega | \hat{p}_i c \hat{\mathcal{L}}_j | \omega \rangle + \frac{\varepsilon^2 \ell^2}{2\tau} \left(\langle \omega | \hat{p}_i \hat{\mathcal{L}}_j \hat{\mathcal{L}}_k | \partial_j \partial_k \omega \rangle + \langle \partial_k \partial_j \omega | \hat{\mathcal{L}}_k \hat{\mathcal{L}}_j \hat{p}_i | \omega \rangle \right) = 0. \quad (84)$$

The zeroth and first momentum equations (82) and (84) are partial differential equations in the matrix elements. The macroscopic equations of motion, a mass continuity equation and a Navier-Stokes equation, come from (82) and (84), respectively.

We can now determine the partial differential equations that describe the dynamics of a quantum lattice gas in local equilibrium. Inserting (78) into the zeroth moment equation (82) and retaining terms up to second order in the smallness gives

$$\begin{aligned} \varepsilon^2 \partial_t \left(m \langle \omega^{(0)} | \omega^{(0)} \rangle \right) + \varepsilon \partial_i \left(m c \langle \omega^{(0)} | \hat{\mathcal{L}}_i | \omega^{(0)} \rangle \right) = \\ - \frac{m \varepsilon^2 \ell^2}{\tau} \partial_i \left(\langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{J}}^{-1} \hat{\mathcal{L}}_j | \partial_j \omega^{(0)} \rangle + \langle \partial_j \omega^{(0)} | \hat{\mathcal{L}}_j (\hat{\mathcal{J}}^{-1})^* \hat{\mathcal{L}}_i | \omega^{(0)} \rangle \right) \\ - \frac{m \varepsilon^2 \ell^2}{2\tau} \left(\langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \partial_i \partial_j \omega^{(0)} \rangle + \langle \partial_i \partial_j \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \omega^{(0)} \rangle \right), \end{aligned} \quad (85)$$

using the fact that $\hat{m} = m \mathbf{1}$. Identifying the local equilibrium mass density (46), $\rho = m \langle \omega^{(0)} | \omega^{(0)} \rangle$, and momentum density (47), $\rho v_i = m c \langle \omega^{(0)} | \hat{\mathcal{L}}_i | \omega^{(0)} \rangle$, we arrive at the hydrodynamic equation for mass flow

$$\begin{aligned} \varepsilon^2 \partial_t \rho + \varepsilon \partial_i (\rho v_i) = \\ - \frac{m \varepsilon^2 \ell^2}{\tau} \partial_i \left(\langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{J}}^{-1} \hat{\mathcal{L}}_j | \partial_j \omega^{(0)} \rangle + \langle \partial_j \omega^{(0)} | \hat{\mathcal{L}}_j (\hat{\mathcal{J}}^{-1})^* \hat{\mathcal{L}}_i | \omega^{(0)} \rangle \right) \\ - \frac{m \varepsilon^2 \ell^2}{2\tau} \left(\langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \partial_i \partial_j \omega^{(0)} \rangle + \langle \partial_i \partial_j \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \omega^{(0)} \rangle \right), \end{aligned} \quad (86)$$

Similarly, by inserting (78) into the first moment equation (84) and retaining terms up to second order in the smallness gives the hydrodynamic equation for momentum flow

$$\begin{aligned} \varepsilon^2 \partial_t (\rho v_i) + \varepsilon \partial_j \Pi_{ij}^{\text{ideal}} = \\ -\varepsilon^2 m c^2 \ell \partial_j \left(\langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j \hat{\mathcal{J}}^{-1} \hat{\mathcal{L}}_k | \partial_k \omega^{(0)} \rangle + \langle \partial_k \omega^{(0)} | \hat{\mathcal{L}}_k (\hat{\mathcal{J}}^{-1})^* \hat{\mathcal{L}}_j \hat{\mathcal{L}}_i | \omega^{(0)} \rangle \right) \\ - \frac{\varepsilon^2}{2} m c^2 \ell \left(\langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j \hat{\mathcal{L}}_k | \partial_j \partial_k \omega^{(0)} \rangle + \langle \partial_k \partial_j \omega^{(0)} | \hat{\mathcal{L}}_k \hat{\mathcal{L}}_j \hat{\mathcal{L}}_i | \omega^{(0)} \rangle \right), \end{aligned} \quad (87)$$

where the ideal part of the momentum flux density tensor is defined as

$$\Pi_{ij}^{\text{ideal}} \equiv m c^2 \langle \omega^{(0)} | \hat{\mathcal{L}}_i \hat{\mathcal{L}}_j | \omega^{(0)} \rangle. \quad (88)$$

To obtain the macroscopic equations of motion, we have to determine the value of the matrix elements appearing in (86), (87) and (88). This is carried out in Appendix B. The result is that in the incompressible limit ($\nabla \cdot \mathbf{v} = 0$), all the matrix elements on the R.H.S. of (86) vanish leaving us with a mass continuity equation

$$\partial_t \rho + \partial_i (\rho v_i) = 0, \quad (89)$$

and the matrix elements on the R.H.S. of (87) do not vanish leaving us with a viscous Navier-Stokes equation¹²

$$\partial_t (\rho v_i) + \partial_g (g \rho v_i v_j) = -\partial_i P + \rho \nu \partial^2 v_i. \quad (90)$$

In (90), the pressure is

$$P = \rho \frac{c^2}{D} \left(1 - g \frac{v^2}{c^2} \right), \quad (91)$$

and the kinematic shear viscosity is

$$\nu = \frac{\ell^2}{\tau(D+2)} \left(\frac{1}{\kappa_\eta} - \frac{1}{2} \right). \quad (92)$$

The form of the kinematic viscosity (92) is identical to that for the classical lattice gas as found by Henon [20]. However, in the case of the quantum lattice gas, the value of κ_η can be different than the classical value because of quantum mechanical interference of outgoing collision possibilities.

11 Summary

In this paper we predicted the macroscopic behavior of a lattice-gas quantum computer. The following observations are made:

¹² As is the case for the classical lattice gas, there is a density dependent prefactor appearing in the convective term and the pressure, $g(d) = \frac{D}{D+2} \frac{1-2d}{1-d}$.

1. The Ψ -space unitary collision matrix, \hat{C} , is successively blocked, first over the on-site 2^B -dimensional manifold, $\hat{C} = \bigotimes_{x=1}^V \hat{U}$. Each block \hat{U} is also block diagonal over the equivalence classes.
2. The projection operator, $\hat{\Gamma}$, periodically causes a partial collapse of the on-site superpositions. So a quantum computer with many qubits can simulate a quantum lattice gas with only short-term and short-range entanglement and coherence of qubits.
3. Streaming of complex amplitudes (or the associated qubits) occurs in analogy to streaming in a classical lattice gas and does not cause global entanglement in the quantum lattice gas because of the application of $\hat{\Gamma}$.
4. The quantum lattice gas can be understood as existing between two limits, a fully coherent quantum system and a classical system. And its single-particle distribution function, $f_a = |\omega_a|^2 = \langle q_a | \hat{n} | q_a \rangle = \langle \psi | \hat{n}_a | \psi \rangle$, has the form $f_a = 1 / (\exp(\alpha\rho + \beta\hat{e}_a \cdot \mathbf{p} + \gamma E) + 1)$.
5. Like the lattice Boltzmann equation approach to simulate fluid dynamics, the quantum lattice gas is a noiseless method that directly codes the particle dynamics at the mesoscopic scale. However, unlike the lattice Boltzmann BGK collision operator, the quantum lattice-gas collision operator obeys detailed balance. Hence, the method is unconditionally stable.
6. The macroscopic hydrodynamic behavior is described by a viscous Navier-Stokes equation.

12 Closing Remarks

To mimic the behavior of other physical systems, quantum lattice gases need many qubits. A first generation quantum computer, with a only two qubits, cannot test the behavior of the quantum lattice gas at the macroscopic scale. However, useful tests could be conducted on a network of these first generation machines to test the practicality of the quantum lattice gas formalism. For example, we could test the reliability and computational speed of a network of quantum computers. It is reasonable to expect that the number of qubits will grow exponentially according to Moore's law as various quantum computer designs are realized over time.

We know from experience with classical lattice gases that even though the underlying microscopic dynamics is reversible, dissipative shear viscosity arises at the macroscopic scale—entropy increases while at the same time information is conserved because of microscopic reversibility. The reason for this is that information, initially stored in the spatial correlations of arrangement of particle occupancies, in time is transferred into high order particle-particle correlations. The same process should occur in a quantum lattice gas. We do not yet know, in a quantum lattice gas setting, if there is a way to block this informational transfer mechanism and thereby reduce dissipation at the macroscopic scale.

Why consider constructing a quantum computer following the lattice gas paradigm, when a general purpose quantum computer could simulate a quantum lattice gas? The answer to this question is three-fold: (1) because any member of

the appropriate unitary group associated with an equivalence class block of the collision operator is sufficient for the recover of Navier-Stokes hydrodynamics at the macroscopic scale, so the lattice-gas quantum computer is robust; (2) short-term coherence among only a small number of nearby qubits is needed; and (3) its behavior is predictable by analytic means.

The quantum lattice gas method can be straightforwardly applied to three-dimensional fluid simulations (the two-dimensional case was treated in this paper because of its simplicity) and also applied to model other physical systems. Lattice gases are a special case of cellular automata where conservations and detailed balance are imposed and an isotropic spatial lattice is used. With these few restrictions removed, the method presented in this paper represents a general computational system called a quantum cellular automaton.

13 Acknowledgements

I would like to thank Dr Bruce Boghosian and Prof Hugh Pendleton for their helpful discussions.

References

1. Wolfram, S.: Cellular automaton fluids 1: Basic theory. *J. of Stat. Phys.*, **45**, No. 3/4 (1986) 471–526
2. Frisch, U., Hasslacher, B., Pomeau, Y.: Lattice-gas automata for the navier-stokes equation. *Phys. Rev. Lett.*, **56**, No. 14 (1986) 1505–1508
3. Yepez, J.: Lattice-gas quantum computation. *Inter. J. Theor. Phys.*, To appear, (1998) Proceeding of the 7th International Conference on the Discrete Simulation of Fluids, University of Oxford.
4. Feynman, R. P.: Simulating physics with computers. *Inter. J. Theor. Phys.*, **21**, No. 6/7 (1982) 467–488
5. Benioff, P.: Quantum mechanical hamiltonian models of turing machines. *J. of Stat. Phys.*, **29**, No. 3 (1982) 515–547
6. Deutsch, D.: Quantum theory, the church-turing principle and the universal quantum computer. *Proc. Roy. Soc., London*, **A400** (1985) 97
7. Margolus, M.: Quantum computation. In Daniel Greenberger, editor, *New Techniques and Ideas in Quantum Measurement Theory*, Annals of the New York Academy **480** (1985) 487–497
8. Bennett, C. H.: Quantum information and computation. *Physics Today*, Oct (1995) 24–30
9. Hey, A.J.G., Allen, R.W., editors: *Feynman Lectures on Computation*. The Advanced Book Program. Addison-Wesley Publishing Company, Inc. (1996)
10. Cory, D.G., Fahmy, A.F., Havel, T.F.: Ensemble quantum computing by nuclear magnetic resonance spectroscopy. Harvard Univ. Center for Research in Comp. Tech., Aiken Comp. Lab., Tech Rep **TR-10-96** (1996)
11. Calderbank, A.R., Shor, P.W.: Good quantum error correcting codes exist. *LANL archive: quant-ph/9512032* (1995)
12. Seth Lloyd.: Quantum-mechanical computers. *Physics Today*, Oct (1995) 140–145

13. Ekert, A., Jozsa, R.: Quantum computation and shor's factoring algorithm. *Rev. of Mod. Phys.*, **68**, No. 3 (1996) 733–753
14. Fredkin, E., Toffoli, T.: Conservative logic. *Inter. J. Theor. Phys.*, **21**, No. 3/4 (1982) 219–253
15. Bennett, C.H.: Logical reversibility of computation. *IBM J. Res. Dev.*, **6** (1979) 525–532
16. Bennett, C.H.: Thermodynamics of computation—a review. *Inter. J. of Theor. Phys.*, **21** (1982) 219–253
17. Sauro Succì.: Lattice boltzmann equation for quantum mechanics. *Physica D*, **69** (1993) 327–332
18. Boghosian, B.M., Taylor, W.: A quantum lattice-gas model for the many-particle schrodinger equation in d-dimensions. *Phys. Rev. E*, **57**, No. 1 (1998) 54–66
19. Yepez, J.: Lattice gas for superfluid helium ii. *Inter. J. Theor. Phys.* Accepted 1996. To Appear (1998). Presented at 6th Inter. Conf. on Discrete Models for Fluid Mech., Boston Univ. Center for Comp. Sci, Aug 1996.
20. Hénon, M.: Viscosity of a lattice gas. In Gary D. Doolean, editor, *Lattice Gas Methods for Partial Differential Equations*, Santa Fe Institute, Addison-Wesley Publishing Company (1990) 179–207
21. Frisch, U., d'Humières, D., Hasslacher, B., Lallemand, P., Pomeau, Y., Rivet, J.P.: Lattice gas hydrodynamics in two and three dimensions. *Comp. Sys.*, **1** (1987) 649–707
22. Yepez, J.: Lattice gas dynamics: Volume 1 viscous fluids. Technical Report **PL-TR-96-2122(I)**, Air Force Research Laboratory, AFRL/VSBE Hanscom AFB, MA, Nov (1996)

A Single-Particle Distribution Function

The single-particle distribution function has the form

$$f(z_a) = \frac{1}{z_a + 1}, \quad (93)$$

where the natural log of the *fugacity*

$$\ln z_a = \alpha \rho + \beta \hat{e}_a \cdot \mathbf{p} + \gamma E \quad (94)$$

is a linear combination of the conserved scalar quantities, the mass ρ , the momentum component $\hat{e}_a \cdot \mathbf{p}$ along the lattice direction \hat{e}_a , and the energy E at a lattice site. The real numbered coefficients α , β , and γ are free parameters that we will determine. It is convenient to define the momentum and energy independent part of the fugacity as

$$z_o \equiv e^{\alpha \rho}. \quad (95)$$

Since $f_a(z_o) = d$ is the reduced density, $d \equiv \frac{\rho}{mB}$, we must set

$$z_o = \frac{1 - d}{d}. \quad (96)$$

This fixes the coefficient α . To fix the coefficients β and γ , we can specify two moments of the single-particle distribution function as constraint conditions. We begin by Taylor expanding the single-particle distribution function $f(z_a)$ about z_o

$$f(z_a) = d + f'(z_o)\delta z + \frac{1}{2}f''(z_o)(\delta z)^2 + \dots \quad (97)$$

The derivative of f evaluated at z_o are

$$f'(z) = \frac{-1}{(z+1)^2} \longrightarrow f'(z_o) = -d^2 \quad (98)$$

and

$$f''(z) = \frac{2}{(z+1)^3} \longrightarrow f''(z_o) = 2d^3, \quad (99)$$

so

$$f(z_a) \cong d [1 - d\delta z + d^2(\delta z)^2]. \quad (100)$$

To determine δz , we begin by writing the fugacity in series form

$$z_a = z_o \left[\sum_{k=0}^{\infty} \frac{(\beta \hat{e}_a \cdot \mathbf{p})^k}{k!} \right] \left[\sum_{k=0}^{\infty} \frac{(\gamma E)^k}{k!} \right]. \quad (101)$$

In the subsonic limit, $\mathbf{p} \ll mc$, keeping terms only to second order in the velocity, the fugacity becomes

$$z_a = z_o \left[1 + \beta \hat{e}_a \cdot \mathbf{p} + \frac{1}{2}(\beta \hat{e}_a \cdot \mathbf{p})^2 \right] (1 + \gamma E) + \mathcal{O}(v^3). \quad (102)$$

since $p \sim v$ and $E \sim v^2$. Then to second order in the velocity, the change in z_a is

$$\delta z_a \equiv z_a - z_o = \left(\frac{1-d}{d} \right) \left[\beta \hat{e}_a \cdot \mathbf{p} + \frac{1}{2}(\beta \hat{e}_a \cdot \mathbf{p})^2 + \gamma E \right] + \mathcal{O}(v^3) \quad (103)$$

and the square of the change is

$$(\delta z_a)^2 = \left(\frac{1-d}{d} \right)^2 \beta^2 (\hat{e}_a \cdot \mathbf{p})^2 + \mathcal{O}(v^3). \quad (104)$$

Inserting the expressions for δz and $(\delta z)^2$ into the Taylor expansion of $f(z_a)$ we have

$$\begin{aligned} f(z_a) &= d \left\{ 1 - (1-d) \left[\beta \hat{e}_a \cdot \mathbf{p} + \frac{1}{2}(\beta \hat{e}_a \cdot \mathbf{p})^2 + \gamma E \right] + (1-d)^2 (\hat{e}_a \cdot \mathbf{p})^2 \right\} \\ &= d \left[1 - (1-d) (\beta \hat{e}_a \cdot \mathbf{p} + \gamma E) + \frac{1}{2}(1-d)(1-2d)\beta^2 (\hat{e}_a \cdot \mathbf{p})^2 \right]. \end{aligned} \quad (105)$$

We have the freedom to choose the coefficients β and γ to parameterized the distribution function as we see fit to satisfy any two constraints. Consider a

parameterization that fixes the value of the coefficients β and γ by using the following moments for the mass density and momentum density

$$\rho = m \sum_{a=1}^B f_a \quad (106)$$

$$\rho \mathbf{v} = mc \sum_{a=1}^B \hat{e}_a f_a. \quad (107)$$

The parameterization may be termed the *non-Galilean parametrization*. Constraints (106) and (107) are typically used in the formulation of classical lattice gases. The single particle distribution function using this non-Galilean parameterization was first found in the mid 1980's by the US researchers Wolfram and Hasslacher and by the French researchers Frisch, d'Humières, Lallemand, Pomeau, and Rivet [121]. Their derivation of (112) given below is different then the derivation presented in this section; they used only two free coefficients in the expression for the fugacity, one for the mass and the other for the momentum; whereas we use three free coefficients. The reason for using only two free parameters is that in the standard single-speed classical lattice-gas construction, the energy is degenerate with the mass, so it was deemed unnecessary to keep a separate free coefficient for the energy. Using (106) and (107) as constraint equations gives us a non-unity density-dependent prefactor in the convective term in the hydrodynamic flow equation.

Inserting (105) into (107), the odd term in the distribution function expansion survives the first moment sum over lattice directions; the odd term is the one linear in the momentum. This fixes the value of β to be

$$\beta = -\frac{D}{1-d} \quad (108)$$

so the distribution function becomes

$$f_a = d \left[1 + D \hat{e}_a \cdot \mathbf{p} + \frac{D^2}{2} \frac{1-2d}{1-d} (\hat{e}_a \cdot \mathbf{p})^2 + (1-d)\gamma E \right]. \quad (109)$$

Inserting (109) into (106), all the even terms that survive the sum over lattice directions must add to zero. This fixes the value of γ as follows

$$\frac{D}{2} \frac{1-2d}{1-d} p^2 - (1-d)\gamma E = 0 \quad (110)$$

or

$$\gamma E = D \frac{1-2d}{(1-d)^2} \frac{p^2}{2}. \quad (111)$$

Therefore, the non-Galilean parameterized distribution function is

$$f_a = d \left[1 + D e_{ai} p_i + \frac{D(D+2)}{2} g(d) Q_{aij} p_i p_j \right], \quad (112)$$

where the density dependent prefactor $g(d)$ is defined

$$g(d) \equiv \frac{D}{D+2} \frac{1-2d}{1-d} \quad (113)$$

and the traceless second-rank tensor \hat{Q}_a is defined

$$Q_{aij} \equiv e_{ai}e_{aj} - \frac{\delta_{ij}}{D}. \quad (114)$$

\hat{Q}_a is an isotropic symmetric tensor. This mass-energy degeneracy leads to an anomalous description of the lattice-gas fluid's behavior. Let us see why. The second moment of (112) gives the momentum flux density

$$mc^2 \sum_{a=1}^B e_{ai}e_{aj}f_a = P\delta_{ij} + g\rho v_i v_j. \quad (115)$$

The density-dependent prefactor g appears in the nonlinear convective term, so this parametrization does indeed give rise to non-Galilean fluid flow. The pressure in (115) has a spurious quadratic velocity dependence

$$P = \rho c_s^2 \left(1 - g \frac{v^2}{c^2} \right). \quad (116)$$

B Determination of the Matrix Elements

For the triangular lattice, tensors made up of products of the lattice vectors are symmetric and isotropic [11]. We have the following identities

$$\langle \mathbf{1} | \mathcal{L}_i | \mathbf{1} \rangle = 0 \quad (117)$$

$$\langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle = \frac{B}{D} \delta_{ij} \quad (118)$$

$$\langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k | \mathbf{1} \rangle = 0 \quad (119)$$

$$\langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k \mathcal{L}_l | \mathbf{1} \rangle = \frac{B}{D(D+2)} (\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \quad (120)$$

The Jacobian of the collision operator is circulant. Its eigenvectors corresponding to the nonzero eigenvalues span the *kinetic space*, which contains a *viscous subspace* characterized by a degenerate eigenvalue, denoted by κ_η [22]. The eigenvectors in the viscous subspace are $\mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle$, for $i \neq j$. Therefore, we have

$$\mathcal{J} \mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle = \kappa_\eta \mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle, \quad (121)$$

or inverting this over the kinetic viscous modes

$$\mathcal{J}^{-1} \mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle = \frac{1}{\kappa_\eta} \mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle. \quad (122)$$

Using these identities along with the epsilon expansion of $|\omega\rangle$, we can work out the value of the matrix elements which appear in the mass and momentum hydrodynamic equations of the quantum lattice gas at the macroscopic scale. We have

$$\begin{aligned}
\langle \omega^{(0)} | \mathcal{L}_i \mathcal{L}_j | \omega^{(0)} \rangle &= d \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j | \mathbf{1} \rangle + \frac{dD}{c} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k | \mathbf{1} \rangle v_k + \\
&\frac{gdD(D+2)}{2c^2} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \left(\mathcal{L}_k \mathcal{L}_l - \frac{\delta_{kl}}{D} \right) | \mathbf{1} \rangle v_k v_l \\
&= \frac{dB}{D} \delta_{ij} + \frac{gdB}{2c^2} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) v_k v_l - \frac{gd(D+2)B}{2Dc^2} \delta_{ij} v^2 \\
&= \frac{dB}{D} \left(1 - g \frac{v^2}{c^2} \right) \delta_{ij} + gdB \frac{v_i v_j}{c^2}. \tag{123}
\end{aligned}$$

$$\begin{aligned}
\langle \omega^{(0)} | \mathcal{L}_i \mathcal{L}_j | \partial_i \partial_j \omega^{(0)} \rangle &= \frac{dD}{2c} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k | \mathbf{1} \rangle \partial_i \partial_j v_k \\
&= 0. \tag{124}
\end{aligned}$$

$$\begin{aligned}
\langle \omega^{(0)} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k | \partial_j \partial_k \omega^{(0)} \rangle &= \frac{dD}{2c} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k \mathcal{L}_l | \mathbf{1} \rangle \partial_j \partial_k v_l \\
&= \frac{dB}{2c(D+2)} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \partial_j \partial_k v_l \\
&= \frac{dB}{2c(D+2)} (2\partial_i \partial_k v_k + \partial^2 v_i). \tag{125}
\end{aligned}$$

$$\begin{aligned}
\langle \omega^{(0)} | \mathcal{L}_i \mathcal{J}^{-1} \mathcal{L}_j | \partial_j \omega^{(0)} \rangle &= \frac{dD}{2c} \langle \mathbf{1} | \mathcal{L}_i \mathcal{J}^{-1} \mathcal{L}_j \mathcal{L}_k | \mathbf{1} \rangle \partial_j v_k \\
&= \frac{dD}{2c\kappa_\eta} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k | \mathbf{1} \rangle \partial_j v_k \\
&= 0. \tag{126}
\end{aligned}$$

$$\begin{aligned}
\langle \omega^{(0)} | \mathcal{L}_i \mathcal{L}_j \mathcal{J}^{-1} \mathcal{L}_k | \partial_k \omega^{(0)} \rangle &= \frac{dD}{2c} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{J}^{-1} \mathcal{L}_k \mathcal{L}_l | \mathbf{1} \rangle \partial_k v_l \\
&= \frac{dD}{2c\kappa_\eta} \langle \mathbf{1} | \mathcal{L}_i \mathcal{L}_j \mathcal{L}_k \mathcal{L}_l | \mathbf{1} \rangle \partial_k v_l \\
&= \frac{dB}{2c\kappa_\eta(D+2)} (\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \partial_k v_l \\
&= \frac{dB}{2c\kappa_\eta(D+2)} (\partial_k v_k \delta_{ij} + \partial_i v_j + \partial_j v_i). \tag{127}
\end{aligned}$$

Quantum Entanglement and the Communication Complexity of the Inner Product Function

Richard Cleve^{1*}, Wim van Dam², Michael Nielsen³, and Alain Tapp^{4**}

¹ University of Calgary

Department of Computer Science, University of Calgary, Calgary, Alberta, Canada
T2N 1N4. cleve@cpsc.ucalgary.ca.

² University of Oxford and CWI, Amsterdam

Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road,
Oxford OX1 3PU, U.K. wimvdam@mildred.physics.ox.ac.uk.

³ Los Alamos National Laboratory and University of New Mexico

T-6 Theoretical Astrophysics, Los Alamos National Laboratory, U.S.A.
mnielsen@tangelo.phys.unm.edu.

⁴ Université de Montréal

Département IRO, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada
H3C 3J7. tappa@iro.umontreal.ca.

Abstract. We consider the communication complexity of the binary inner product function in a variation of the two-party scenario where the parties have an *a priori* supply of particles in an entangled quantum state. We prove linear lower bounds for both exact protocols, as well as for protocols that determine the answer with bounded-error probability. Our proofs employ a novel kind of “quantum” reduction from a quantum information theory problem to the problem of computing the inner product. The communication required for the former problem can then be bounded by an application of Holevo’s theorem. We also give a specific example of a probabilistic scenario where entanglement reduces the communication complexity of the inner product function by one bit.

1 Introduction and Summary of Results

The *communication complexity* of a function $f : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$ is defined as the minimum amount of communication necessary among two parties, conventionally referred to as Alice and Bob, in order for, say, Bob to acquire the value of $f(x, y)$, where, initially, Alice is given x and Bob is given y . This scenario was introduced by Yao [16] and has been widely studied (see [13] for a survey). There are a number of technical choices in the model, such as: whether the communication cost is taken as the worst-case (x, y) , or the average-case (x, y) with respect to some probability distribution; whether the protocols are

* Research initiated while visiting the Université de Montréal and supported in part by Canada’s NSERC.

** Research supported in part by Canada’s NSERC.

deterministic or probabilistic (and, for probabilistic protocols, whether the parties have independent random sources or a shared random source); and, what correctness probability is required.

The communication complexity of the *inner product modulo two (IP)* function

$$IP(x, y) = x_1y_1 + x_2y_2 + \cdots + x_ny_n \bmod 2 \quad (1)$$

is fairly well understood in the above “classical” models. For worst-case inputs and deterministic errorless protocols, the communication complexity is n and, for randomized protocols (with either an independent or a shared random source), uniformly distributed or worst-case inputs, and with error probability $\frac{1}{2} - \delta$ required, the communication complexity is $n - O(\log(1/\delta))$ [7] (see also [13]).

In 1993, Yao [17] introduced a variation of the above classical communication complexity scenarios, where the parties communicate with *qubits*, rather than with bits. Protocols in this model are at least as powerful as probabilistic protocols with independent random sources. Kremer [12] showed that, in this model, the communication complexity of *IP* is $\Omega(n)$, whenever the required correctness probability is $1 - \varepsilon$ for a constant $0 \leq \varepsilon < \frac{1}{2}$ (Kremer attributes the proof methodology to Yao).

Cleve and Buhrman [8] (see also [6]) introduced another variation of the classical communication complexity scenario that also involves quantum information, but in a different way. In this model, Alice and Bob have an initial supply of particles in an entangled quantum state, such as Einstein-Podolsky-Rosen (EPR) pairs, but the communication is still in terms of classical bits. They showed that the entanglement enables the communication for a specific problem to be reduced by one bit. Any protocol in Yao’s qubit model can be simulated by a protocol in this entanglement model with at most a factor two increase in communication: each qubit can be “teleported” [3] by sending two classical bits in conjunction with an EPR pair of entanglement. On the other hand, we are aware of no similar simulation of protocols in the entanglement model by protocols in the qubit model, and, thus, the entanglement model is potentially stronger.

In this paper, we consider the communication complexity of *IP* in two scenarios: with prior entanglement and qubit communication; and with prior entanglement and classical bit communication. As far as we know, the proof methodology of the lower bound in the qubit communication model without prior entanglement [12] does not carry over to either of these two models. Nevertheless, we show $\Omega(n)$ lower bounds in these models.

To state our lower bounds more precisely, we introduce the following notation. Let $f : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$ be a communication problem, and $0 \leq \varepsilon < \frac{1}{2}$. Let $Q_\varepsilon^*(f)$ denote the communication complexity of f in terms of *qubits*, where quantum entanglement is available and the requirement is that Bob determines the correct answer with probability at least $1 - \varepsilon$ (the $*$ superscript is intended to highlight the fact that prior entanglement is available). Also, let $C_\varepsilon^*(f)$ denote the corresponding communication complexity of f in the scenario where the

communication is in terms of *bits* (again, quantum entanglement is available and Bob is required to determine the correct answer with probability at least $1 - \varepsilon$). When $\varepsilon = 0$, we refer to the protocols as *exact*, and, when $\varepsilon > 0$, we refer to them as *bounded-error* protocols. With this notation, our results are:

$$Q_0^*(IP) = \lceil n/2 \rceil \quad (2)$$

$$Q_\varepsilon^*(IP) \geq \frac{1}{2}(1 - 2\varepsilon)^2 n - \frac{1}{2} \quad (3)$$

$$C_0^*(IP) = n \quad (4)$$

$$C_\varepsilon^*(IP) \geq \max(\frac{1}{2}(1 - 2\varepsilon)^2, (1 - 2\varepsilon)^4) n - \frac{1}{2} \quad (5)$$

Note that all the lower bounds are $\Omega(n)$ whenever ε is held constant. Also, these results subsume the lower bounds in [12], since the qubit model defined by Yao [17] differs from the bounded-error qubit model defined above only in that it does not permit a prior entanglement.

Our lower bound proofs employ a novel kind of “quantum” reduction between protocols, which reduces the problem of communicating, say, n bits of information to the IP problem. It is noteworthy that, in classical terms, it can be shown that there is no such reduction between the two problems. The appropriate cost associated with communicating n bits is then lower-bounded by the following nonstandard consequence of Holevo’s theorem.

Theorem 1: *In order for Alice to convey n bits of information to Bob, where quantum entanglement is available and qubit communication in either direction is permitted, Alice must send Bob at least $\lceil n/2 \rceil$ qubits. This holds regardless of the prior entanglement and the qubit communication from Bob to Alice. More generally, for Bob to obtain m bits of mutual information with respect to Alice’s n bits, Alice must send at least $\lceil m/2 \rceil$ qubits.*

A slight generalization of Theorem 1 is described and proven in the Appendix.

It should be noted that, since quantum information subsumes classical information, our results also represent new proofs of nontrivial lower bounds on the *classical* communication complexity of IP , and our methodology is fundamentally different from those previously used for classical lower bounds.

Finally, with respect to the question of whether quantum entanglement can *ever* be advantageous for protocols computing IP , we present a curious probabilistic scenario with $n = 2$ where prior entanglement enables one bit of communication to be saved.

2 Bounds for Exact Qubit Protocols

In this section, we consider exact qubit protocols computing IP , and prove Eq. (2). Note that the upper bound follows from so-called “superdense coding” [4]: by sending $\lceil n/2 \rceil$ qubits in conjunction with $\lceil n/2 \rceil$ EPR pairs, Alice can transmit her n classical bits of input to Bob, enabling him to evaluate IP . For the lower bound, we consider an arbitrary exact qubit protocol that computes IP , and convert it (in two stages) to a protocol for which Theorem 1 applies.

For convenience, we use the following notation. If an m -qubit protocol consists of m_1 qubits from Alice to Bob and m_2 qubits from Bob to Alice then we refer to the protocol as an (m_1, m_2) -qubit protocol.

2.1 Converting Exact Protocols into Clean Form

A *clean protocol* is a special kind of qubit protocol that follows the general spirit of the reversible programming paradigm in a quantum setting. Namely, one in which all qubits incur no net change, except for one, which contains the answer.

In general, the initial state of a qubit protocol is of the form

$$\underbrace{|y_1, \dots, y_n\rangle|0, \dots, 0\rangle}_{\text{Bob's qubits}} \underbrace{|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle}_{\text{Alice's qubits}}, \quad (6)$$

where $|\Phi_{BA}\rangle$ is the state of the entangled qubits shared by Alice and Bob, and the $|0, \dots, 0\rangle$ states can be regarded as “ancillas”. At each turn, a player performs some transformation (which, without loss of generality, can be assumed to be unitary) on all the qubits in his/her possession and then sends a subset of these qubits to the other player. Note that, due to the communication, the qubits possessed by each player varies during the execution of the protocol. At the end of the protocol, Bob measures one of his qubits which is designated as his *output*.

We say that a protocol which exactly computes a function $f(x, y)$ is *clean* if, when executed on the initial state

$$|z\rangle|y_1, \dots, y_n\rangle|0, \dots, 0\rangle|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle, \quad (7)$$

results in the final state

$$|z + f(x, y)\rangle|y_1, \dots, y_n\rangle|0, \dots, 0\rangle|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle \quad (8)$$

(where the addition is mod 2). The “input”, the ancilla, and initial entangled qubits will typically change states during the execution of the protocol, but they are reset to their initial values at the end of the protocol.

It is straightforward to transform an exact (m_1, m_2) -qubit protocol into a clean $(m_1 + m_2, m_1 + m_2)$ -qubit protocol that computes the same function. To reset the bits of the input, the ancilla, and the initial entanglement, the protocol is run once, except the output is not measured, but recorded and then the protocol is run in the *backwards* direction to “undo the effects of the computation”. The output is recorded on a *new* qubit of Bob (with initial state $|z\rangle$) which is control-negated with the output qubit of Bob (that is in the state $|f(x, y)\rangle$) as the control. Note that, for each qubit that Bob sends to Alice when the protocol is run forwards, Alice sends the qubit to Bob when run in the backwards direction. Running the protocol backwards resets all the qubits—except Bob’s new one—to their original states. The result is an $(m_1 + m_2, m_1 + m_2)$ -qubit protocol that maps state $\textcircled{7}$ to state $\textcircled{8}$.

2.2 Reduction from Communication Problems

We now show how to transform a clean $(m_1 + m_2, m_1 + m_2)$ -qubit protocol that exactly computes IP for inputs of size n , to an $(m_1 + m_2, m_1 + m_2)$ -qubit protocol that transmits n bits of information from Alice to Bob. This is accomplished in four stages:

1. Bob initializes his qubits indicated in Eq. (7) with $z = 1$ and $y_1 = \dots = y_n = 0$.
2. Bob performs a Hadamard transformation on each of his first $n + 1$ qubits.
3. Alice and Bob execute the clean protocol for the inner product function.
4. Bob again performs a Hadamard transformation on each of his first $n + 1$ qubits.

Let $|B_i\rangle$ denote the state of Bob's first $n + 1$ qubits after the i^{th} stage. Then

$$|B_1\rangle = |1\rangle|0, \dots, 0\rangle \quad (9)$$

$$|B_2\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{a, b_1, \dots, b_n \in \{0,1\}} (-1)^a |a\rangle |b_1, \dots, b_n\rangle \quad (10)$$

$$\begin{aligned} |B_3\rangle &= \frac{1}{\sqrt{2^{n+1}}} \sum_{a, b_1, \dots, b_n \in \{0,1\}} (-1)^a |a + b_1 x_1 + \dots + b_n x_n\rangle |b_1, \dots, b_n\rangle \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{c, b_1, \dots, b_n \in \{0,1\}} (-1)^{c + b_1 x_1 + \dots + b_n x_n} |c\rangle |b_1, \dots, b_n\rangle \end{aligned} \quad (11)$$

$$|B_4\rangle = |1\rangle |x_1, \dots, x_n\rangle, \quad (12)$$

where, in Eq. (11), the substitution $c = a + b_1 x_1 + \dots + b_n x_n$ has been made (and arithmetic over bits is taken mod 2). The above transformation was inspired by the reading of [14] (see also [5]).

Since the above protocol conveys n bits of information (namely, x_1, \dots, x_n) from Alice to Bob, by Theorem 1, we have $m_1 + m_2 \geq n/2$. Since this protocol can be constructed from an arbitrary exact (m_1, m_2) -qubit protocol for IP , this establishes the lower bound of Eq. (2).

Note that, classically, no such reduction is possible. For example, if a clean protocol for IP is executed in any classical context, it can never yield more than one bit of information to Bob (whereas, in this quantum context, it yields n bits of information to Bob).

3 Lower Bounds for Bounded-Error Qubit Protocols

In this section we consider bounded-error qubit protocols for IP , and prove Eq. (3). Assume that some qubit protocol P computes IP correctly with probability at least $1 - \varepsilon$, where $0 < \varepsilon < \frac{1}{2}$. Since P is not exact, the constructions from the previous section do not work exactly. We analyze the extent by which they err.

First, the construction of Section 2.1 will not produce a protocol in clean form; however, it will result in a protocol which *approximates* an exact clean

protocol (this type of construction was previously carried out in a different context by Bennett *et al.* [2]).

Denote the initial state as

$$|y_1, \dots, y_n\rangle|0, \dots, 0\rangle|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle. \quad (13)$$

Also, assume that, in protocol P , Bob never changes the state of his input qubits $|y_1, \dots, y_n\rangle$ (so the first n qubits never change). This is always possible, since he can copy y_1, \dots, y_n into his ancilla qubits at the beginning. After executing P until just before the measurement occurs, the state of the qubits must be of the form

$$\alpha|y_1, \dots, y_n\rangle|x \cdot y\rangle|J\rangle + \beta|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle, \quad (14)$$

where $|\alpha|^2 \geq (1 - \varepsilon)$ and $|\beta|^2 \leq \varepsilon$. In the above, the $n + 1^{\text{st}}$ qubit is the designated output, $x \cdot y$ denotes the inner product of x and y , and $\overline{x \cdot y}$ denotes the negation of this inner product. In general, α , β , $|J\rangle$, and $|K\rangle$ may depend on x and y .

Now, suppose that the procedure described in Section 2.1 for producing a clean protocol in the exact case is carried out for P . Since, in general, the answer qubit is not in the state $|x \cdot y\rangle$ —or even in a pure basis state—this does not produce the final state

$$|z + x \cdot y\rangle|y_1, \dots, y_n\rangle|0, \dots, 0\rangle|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle. \quad (15)$$

However, let us consider the state that is produced instead. After introducing the *new* qubit, initialized in basis state $|z\rangle$, and applying P , the state is

$$|z\rangle(\alpha|y_1, \dots, y_n\rangle|x \cdot y\rangle|J\rangle + \beta|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle). \quad (16)$$

After applying the controlled-NOT gate, the state is

$$\begin{aligned} & \alpha|z + x \cdot y\rangle|y_1, \dots, y_n\rangle|x \cdot y\rangle|J\rangle + \beta|z + \overline{x \cdot y}\rangle|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle \\ &= \alpha|z + x \cdot y\rangle|y_1, \dots, y_n\rangle|x \cdot y\rangle|J\rangle + \beta|z + x \cdot y\rangle|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle \\ & \quad - \beta|z + x \cdot y\rangle|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle + \beta|z + \overline{x \cdot y}\rangle|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle \\ &= |z + x \cdot y\rangle(\alpha|y_1, \dots, y_n\rangle|x \cdot y\rangle|J\rangle + \beta|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle) \\ & \quad + \sqrt{2}\beta\left(\frac{1}{\sqrt{2}}|z + \overline{x \cdot y}\rangle - \frac{1}{\sqrt{2}}|z + x \cdot y\rangle\right)|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle. \end{aligned} \quad (17)$$

Finally, after applying P in reverse to this state, the final state is

$$|z + x \cdot y\rangle|y_1, \dots, y_n\rangle|0, \dots, 0\rangle|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle + \sqrt{2}\beta|M_{x,y,z}\rangle, \quad (18)$$

where

$$|M_{x,y,z}\rangle = \left(\frac{1}{\sqrt{2}}|z + \overline{x \cdot y}\rangle - \frac{1}{\sqrt{2}}|z + x \cdot y\rangle\right)P^\dagger|y_1, \dots, y_n\rangle|\overline{x \cdot y}\rangle|K\rangle. \quad (19)$$

Note that the vector $\sqrt{2}\beta|M_{x,y,z}\rangle$ is the difference between what an exact protocol would produce (state (15)) and what is obtained by using the inexact (probabilistic) protocol P (state (18)). There are some useful properties of the

$|M_{x,y,z}\rangle$ states. First, as $y \in \{0,1\}^n$ varies, the states $|M_{x,y,z}\rangle$ are orthonormal, since $|y_1, \dots, y_n\rangle$ is a factor in each such state (this is where the fact that Bob does not change his input qubits is used). Also, $|M_{x,y,0}\rangle = -|M_{x,y,1}\rangle$, since only the $(\frac{1}{\sqrt{2}}|z + \bar{x} \cdot \bar{y}\rangle - \frac{1}{\sqrt{2}}|z + x \cdot y\rangle)$ factor in each such state depends on z .

Call the above protocol \tilde{P} . Now, apply the four stage reduction in Section 2.2, with \tilde{P} in place of an exact clean protocol. The *difference* between the state produced by using \tilde{P} and using an exact clean protocol first occurs after the third stage and is

$$\begin{aligned} & \frac{1}{\sqrt{2^{n+1}}} \sum_{y_1, \dots, y_n, z \in \{0,1\}} (-1)^z \sqrt{2} \beta_y |M_{x,y,z}\rangle \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{y_1, \dots, y_n \in \{0,1\}} \sqrt{2} \beta_y (|M_{x,y,0}\rangle - |M_{x,y,1}\rangle) \\ &= \frac{2}{\sqrt{2^n}} \sum_{y_1, \dots, y_n \in \{0,1\}} \beta_y |M_{x,y,0}\rangle, \end{aligned} \quad (20)$$

which has magnitude bounded above by $2\sqrt{\varepsilon}$, since, for each $y \in \{0,1\}^n$, $|\beta_y|^2 \leq \varepsilon$, and the $|M_{x,y,0}\rangle$ states are orthonormal. Also, the magnitude of this difference does not change when the Hadamard transform in the fourth stage is applied. Thus, the final state is within Euclidean distance $2\sqrt{\varepsilon}$ from

$$|1\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle|\Phi_{BA}\rangle|x_1, \dots, x_n\rangle|0, \dots, 0\rangle. \quad (21)$$

Consider the angle θ between this final state and (21). It satisfies $\sin^2 \theta + (1 - \cos \theta)^2 \leq 4\varepsilon$, from which it follows that $\cos \theta \geq 1 - 2\varepsilon$. Therefore, if Bob measures his first $n+1$ qubits in the standard basis, the probability of obtaining $|1, x_1, \dots, x_n\rangle$ is $\cos^2 \theta \geq (1 - 2\varepsilon)^2$.

Now, suppose that x_1, \dots, x_n are uniformly distributed. Then Fano's inequality (see, for example, [9]) implies that Bob's measurement causes his uncertainty about x_1, \dots, x_n to drop from n bits to less than $(1 - (1 - 2\varepsilon)^2)n + h((1 - 2\varepsilon)^2)$ bits, where $h(x) = -x \log x - (1 - x) \log(1 - x)$ is the binary entropy function. Thus, the mutual information between the result of Bob's measurement and (x_1, \dots, x_n) is at least $(1 - 2\varepsilon)^2 n - h((1 - 2\varepsilon)^2) \geq (1 - 2\varepsilon)^2 n - 1$ bits. By Theorem 1, the communication from Alice to Bob is at least $\frac{1}{2}(1 - 2\varepsilon)^2 n - \frac{1}{2}$ qubits, which establishes Eq. (3).

4 Lower Bounds for Bit Protocols

In this section, we consider exact and bounded-error bit protocols for IP , and prove Eqs. (4) and (5).

Recall that any m -qubit protocol can be simulated by a $2m$ -bit protocol using teleportation [3] (employing EPR pairs of entanglement). Also, if the communication pattern in an m -bit protocol is such that an even number of bits is always sent during each party's turn then it can be simulated by an $m/2$ -qubit protocol by superdense coding [4] (which also employs EPR pairs). However, this

latter simulation technique cannot, in general, be applied directly, especially for protocols where the parties take turns sending single bits.

We can nevertheless obtain a slightly weaker simulation of bit protocols by qubit protocols for IP that is sufficient for our purposes. The result is that, given any m -bit protocol for IP_n (that is, IP instances of size n), one can construct an m -qubit protocol for IP_{2n} . This is accomplished by interleaving two executions of the bit protocol for IP_n to compute two independent instances of inner products of size n . We make two observations. First, by taking the sum (mod 2) of the two results, one obtains an inner product of size $2n$. Second, due to the interleaving, an even number of bits is sent at each turn, so that the above superdense coding technique can be applied, yielding a $(2m)/2 = m$ -qubit protocol for IP_{2n} . Now, Eq. (2) implies $m \geq n$, which establishes the lower bound of Eq. (4) (and the upper bound is trivial).

If the same technique is applied to any m -bit protocol computing IP_n with probability $1 - \varepsilon$, one obtains an m -qubit protocol that computes IP_{2n} with probability $(1 - \varepsilon)^2 + \varepsilon^2 = 1 - 2\varepsilon(1 - \varepsilon)$. Applying Eq. (3) here, with $2n$ replacing n and $2\varepsilon(1 - \varepsilon)$ replacing ε , yields $m \geq (1 - 2\varepsilon)^4 n - \frac{1}{2}$. For $\varepsilon > \frac{2-\sqrt{2}}{4} = 0.146\dots$, a better bound is obtained by simply noting that $C_\varepsilon^* \geq Q_\varepsilon^*$ (since qubits can always be used in place of bits), and applying Eq. (3). This establishes Eq. (5).

5 An Instance Where Prior Entanglement Is Beneficial

Here we will show that in spite of the preceding results, it is still possible that a protocol which uses prior entanglement outperforms all possible classical protocols. This improvement is done in the probabilistic sense where we look at the number of communication bits required to reach a certain reliability threshold for the IP function. This is done in the following setting.

Both Alice and Bob have a 2 bit vector x_1x_2 and y_1y_2 , for which they want to calculate the inner product modulo 2:

$$f(x, y) = x_1y_1 + x_2y_2 \bmod 2 \quad (22)$$

with a correctness-probability of at least $\frac{4}{5}$. It will be shown that with entanglement Alice and Bob can reach this ratio with 2 bits of communication, whereas without entanglement 3 bits are necessary to obtain this success-ratio.

5.1 A Two-Bit Protocol with Prior Entanglement

Initially Alice and Bob share a joint random coin and an EPR-like pair of qubits Q_A and Q_B :

$$\text{state}(Q_A Q_B) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \quad (23)$$

With these attributes the protocol goes as follows.

First Alice and Bob determine by a joint random coin flip¹ who is going to be the ‘sender’ and the ‘receiver’ in the protocol. (We continue the description of the protocol by assuming that Alice is the sender and that Bob is the receiver.) After this, Alice (the sender) applies the rotation $A_{x_1x_2}$ on her part of the entangled pair and measures this qubit Q_A in the standard basis. The result m_A of this measurement is then sent to Bob (the receiver) who continues the protocol.

If Bob has the input string ‘00’, he knows with certainty that the outcome of the function $f(x, y)$ is zero and hence he concludes the protocol by sending the bit 0 to Alice. Otherwise, Bob performs the rotation $B_{y_1y_2}$ on his part of the entangled pair Q_B and measure it in the standard basis yielding the value m_B . Now Bob finishes the protocol by sending to Alice the bit $m_A + m_B \bmod 2$.

Using the rotations shown below and bearing in mind the randomization process in the beginning of the protocol with the joint coin flip, this will be a protocol that uses only 2 bits of classical communication and that gives the correct value of $f(x, y)$ with a probability of at least $\frac{4}{5}$ for every possible combination of x_1x_2 and y_1y_2 .

The unitary transformations used by the sender in the protocol are:

$$\begin{aligned}
 A_{00} &= \begin{pmatrix} \sqrt{\frac{2}{5}} & -i\sqrt{\frac{3}{5}} \\ -i\sqrt{\frac{3}{5}} & \sqrt{\frac{2}{5}} \end{pmatrix} & A_{01} &= \begin{pmatrix} \sqrt{\frac{4}{5}} & \sqrt{\frac{3}{16}} + i\sqrt{\frac{1}{80}} \\ -\sqrt{\frac{3}{16}} + i\sqrt{\frac{1}{80}} & \sqrt{\frac{4}{5}} \end{pmatrix} \\
 A_{10} &= \begin{pmatrix} \sqrt{\frac{4}{5}} & -\sqrt{\frac{3}{16}} + i\sqrt{\frac{1}{80}} \\ \sqrt{\frac{3}{16}} + i\sqrt{\frac{1}{80}} & \sqrt{\frac{4}{5}} \end{pmatrix} & A_{11} &= \begin{pmatrix} \sqrt{\frac{1}{5}} & i\sqrt{\frac{4}{5}} \\ i\sqrt{\frac{4}{5}} & \sqrt{\frac{1}{5}} \end{pmatrix},
 \end{aligned} \tag{24}$$

whereas the receiver uses one of the three rotations:

$$\begin{aligned}
 B_{01} &= \begin{pmatrix} \sqrt{\frac{3}{5}} & -\frac{1}{2} + i\sqrt{\frac{3}{20}} \\ -\frac{1}{2} - i\sqrt{\frac{3}{20}} & -\sqrt{\frac{3}{5}} \end{pmatrix} & B_{10} &= \begin{pmatrix} \sqrt{\frac{3}{5}} & \frac{1}{2} + i\sqrt{\frac{3}{20}} \\ -\frac{1}{2} + i\sqrt{\frac{3}{20}} & \sqrt{\frac{3}{5}} \end{pmatrix} \\
 B_{11} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
 \end{aligned} \tag{25}$$

The matrices were found by using an optimization program that suggested certain numerical values. A closer examination of these values revealed the above analytical expressions.

¹ Because a joint random coin flip can be simulated with an EPR-pair, we can also assume that Alice and Bob start the protocol with two shared EPR-pairs and no random coins.

5.2 No Two-Bit Classical Probabilistic Protocol Exists

Take the probability distribution π on the input strings x and y , defined by:

$$\pi(x, y) = \begin{cases} 0 & \text{iff } x = 00 \text{ or } y = 00 \\ \frac{1}{9} & \text{iff } x \neq 00 \text{ and } y \neq 00 \end{cases} \quad (26)$$

It is easily verified that for this distribution, every *deterministic* protocol with only two bits of communication will have a correctness ratio of at most $\frac{7}{9}$. Using Theorem 3.20 of [13], this shows that every possible randomized protocol with the same amount of communication will have a success ratio of at most $\frac{7}{9}$. (It can also be shown that this $\frac{7}{9}$ bound is tight but we will omit that proof here.) This implies that in order to reach the requested ration of $\frac{4}{5}$, at least three bits of communication are required if we are not allowed to use any prior entanglement.

5.3 Two Qubits Suffice Without Prior Entanglement

A similar result also holds for qubit protocols without prior entanglement [17]. This can be seen by the fact that after Alice applied the rotation $A_{x_1 x_2}$ and measured her qubit Q_A with the result $m_A = 0$, she knows the state of Bob's qubit Q_B exactly. It is therefore also possible to envision a protocol where the parties assume the measurement outcome $m_A = 0$ (this can be done without loss of generality), and for which Alice simply sends this qubit Q_B to Bob, after which Bob finishes the protocol in the same way as prescribed by the 'prior entanglement'-protocol. The protocol has thus become as follows.

First Alice and Bob decide by a random joint coin flip who is going to be the sender and the receiver in protocol. (Again we assume here that Alice is the sender.) Next, Alice (the sender) sends a qubit $|Q_{x_1 x_2}\rangle$ (according to the input string $x_1 x_2$ of Alice and the table [27]) to the receiver Bob who continues the protocol.

$$\begin{aligned} |Q_{00}\rangle &= \sqrt{\frac{2}{5}}|0\rangle - i\sqrt{\frac{3}{5}}|1\rangle & |Q_{01}\rangle &= \sqrt{\frac{4}{5}}|0\rangle + \left(\sqrt{\frac{3}{16}} + i\sqrt{\frac{1}{80}}\right)|1\rangle \\ |Q_{10}\rangle &= \sqrt{\frac{4}{5}}|0\rangle + \left(-\sqrt{\frac{3}{16}} + i\sqrt{\frac{1}{80}}\right)|1\rangle & |Q_{11}\rangle &= \sqrt{\frac{1}{5}}|0\rangle - i\sqrt{\frac{4}{5}}|1\rangle \end{aligned} \quad (27)$$

If Bob has the input string $y_1 y_2 = 00$, he concludes the protocol by sending a zero bit to Alice. In the other case, Bob applies the rotation $B_{y_1 y_2}$ to the received qubit, measures the qubit in the standard basis, and sends this measurement outcome to Alice as the answer of the protocol. By doing so, the same correctness-probability of $\frac{4}{5}$ is reached for the *IP* function with two qubits of communication, whereas the classical setting requires 3 bits of communication as shown above.

Acknowledgments

We would like to thank Gilles Brassard, Harry Buhrman, Peter Høyer, and Tal Mor for their comments about this research. R.C. would like to thank the Laboratoire d'Informatique Théorique et Quantique, Université de Montréal for their

gracious hospitality while this research was initiated. M.N. thanks the Office of Naval Research (Grant No. N00014-93-1-0116).

References

1. H. Araki and E.H. Lieb, "Entropy inequalities", *Commun. Math. Phys.*, Vol. 18, 1970, pp. 160–170.
2. C.H. Bennett, E. Bernstein, G. Brassard, U. Vazirani, "Strengths and weaknesses of quantum computing", *SIAM J. on Comput.*, Vol. 26, No. 5, 1997, pp. 1510–1523.
3. C.H. Bennett, G. Brassard, C. Crépeau, R. Josza, A. Peres, W.K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels", *Phys. Rev. Lett.*, Vol. 70, 1993, pp. 1895–1899.
4. C.H. Bennett and S.J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states", *Phys. Rev. Lett.*, Vol. 69, No. 20, 1992, pp. 2881–2884.
5. E. Bernstein, U. Vazirani, "Quantum Complexity Theory", *SIAM J. Comput.*, Vol. 26, No. 5, 1997, pp. 1411–1473.
6. H. Buhrman, R. Cleve, and W. van Dam, "Quantum Entanglement and Communication Complexity", preprint available from the LANL quant-ph archive 9705033, 1997.
7. B. Chor and O. Goldreich, "Unbiased bits from weak sources of randomness and probabilistic communication complexity", *SIAM J. on Comput.*, Vol. 17, No. 2, pp. 230–261, 1988.
8. R. Cleve and H. Buhrman, "Substituting quantum entanglement for communication", *Phys. Rev. A*, Vol. 56, No. 2, pp. 1201–1204, 1997.
9. T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.
10. A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be complete?", *Phys. Rev.*, Vol. 47, 1935, pp. 777–780.
11. A.S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel", *Problemy Peredachi Informatsii*, Vol. 9, No. 3, 1973, pp. 3–11. English translation *Problems of Information Transmission*, Vol. 9, 1973, pp. 177–183.
12. I. Kremer, "Quantum Communication", Master's Thesis, The Hebrew University of Jerusalem, 1995.
13. E. Kushilevitz and N. Nisan, *Communication Complexity*, Cambridge University Press, 1997.
14. B.M. Terhal and J.A. Smolin, "Superfast quantum algorithms for coin weighing and binary search problems", preprint available from the LANL quant-ph archive 9705041, 1997.
15. B. Schumacher, M. Westmoreland and W. K. Wootters, "Limitation on the amount of accessible information in a quantum channel", *Phys. Rev. Lett.*, Vol. 76, 1996, pp. 3453–3456.
16. A.C. Yao, "Some complexity questions related to distributed computing", *Proc. of the 11th Ann. ACM Symp. on Theory of Computing*, 1979, pp. 209–213.
17. A.C. Yao, "Quantum circuit complexity", *Proc. of the 34th Ann. IEEE Symp. on Foundations of Computer Science*, 1993, pp. 352–361.

Appendix: Capacity Results for Communication Using Qubits

In this appendix, we present results about the quantum resources required to transmit n classical bits between two parties when two-way communication is available. These results are used in the main text in the proof of the lower bound on the communication complexity of the inner product function, and may also be of independent interest.

Theorem 2: *Suppose that Alice possesses n bits of information, and wants to convey this information to Bob. Suppose that Alice and Bob possess no prior entanglement but qubit communication in either direction is allowed. Let n_{AB} be the number of qubits Alice sends to Bob, and n_{BA} the number of qubits Bob sends to Alice (n_{AB} and n_{BA} are natural numbers). Then, Bob can acquire the n bits if and only if the following inequalities are satisfied:*

$$n_{AB} \geq \lceil n/2 \rceil \quad (28)$$

$$n_{AB} + n_{BA} \geq n. \quad (29)$$

More generally, Bob can acquire m bits of mutual information with respect to Alice's n bits if and only if the above equations hold with m substituted for n .

Note that Theorem 1 follows from Theorem 2 because, if the communication from Bob to Alice is not counted then this can be used to set up an arbitrary entanglement at no cost.

Graphically, the capacity region for the above communication problem is shown in Fig. 1. Note the difference with the classical result for communication with bits, where the capacity region is given by the equation $n_{AB} \geq n$; that is, classically, communication from Bob to Alice does not help.

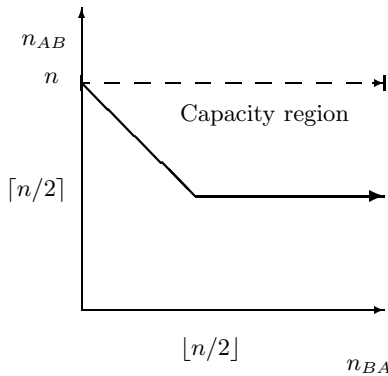


Fig. 1. Capacity region to send n bits from Alice to Bob. n_{AB} is the number of qubits Alice sends to Bob, and n_{BA} is the number of qubits Bob sends to Alice. The dashed line indicates the bottom of the classical capacity region.

Proof of Theorem 2: The sufficiency of Eqns. (28) and (29) follows from the superdense coding technique [4]. The nontrivial case is where $n_{AB} < n$. Bob prepares $n - n_{AB} \leq n_{BA}$ EPR pairs and sends one qubit of each pair to Alice, who can use them in conjunction with sending $n - n_{AB} \leq n_{AB}$ qubits to Bob to transmit $2(n - n_{AB})$ bits to Bob. Alice uses her remaining allotment of $2n_{AB} - n$ qubits to transmit the remaining $2n_{AB} - n$ bits in the obvious way.

The proof that Eqns. (28) and (29) are necessary follows from an application of Holevo's Theorem [11], which we now review. Suppose that a classical information source produces a random variable X . Depending on the value, x , of X , a quantum state with density operator ρ_x is prepared. Suppose that a measurement is made on this quantum state in an effort to determine the value of X . This measurement results in an outcome Y . Holevo's theorem states that the mutual information $I(X : Y)$ between X and Y is bounded by the *Holevo bound* [11]

$$I(X : Y) \leq S(\rho) - \sum_x p_x S(\rho_x), \quad (30)$$

where p_x are the probabilities associated with the different values of X , $\rho = \sum_x p_x \rho_x$, and S is the von Neumann entropy function. The quantity on the right hand side of the Holevo bound is known as the *Holevo chi quantity*, $\chi(\rho_X) = S(\rho) - \sum_x p_x S(\rho_x)$.

Let X be Alice's n bits of information, which is uniformly distributed over $\{0, 1\}^n$. Without loss of generality, it can be assumed that the protocol between Alice and Bob is of the following form. For any value (x_1, \dots, x_n) of X , Alice begins with a set of qubits in state $|x_1, \dots, x_n\rangle|0, \dots, 0\rangle$ and Bob begins with a set of qubits in state $|0, \dots, 0\rangle$. The protocol first consists of a sequence of steps, where at each step one of the following processes takes place.

1. Alice performs a unitary operation on the qubits in her possession.
2. Bob performs a unitary operation on the qubits in his possession.
3. Alice sends a qubit to Bob.
4. Bob sends a qubit to Alice.

After these steps, Bob performs a measurement on the qubits in his possession, which has outcome Y . (Note that one might imagine that the initial states could be mixed and that measurements could be performed in addition to unitary operations; however, these processes can be simulated using standard techniques involving ancilla qubits.)

Let ρ_i^X be the density operator of the set of qubits that are in Bob's possession after i steps have been executed. Due to Holevo's Theorem, it suffices to upper bound the final value of $\chi(\rho_i^X)$ —which is also bounded above by $S(\rho_i)$. We consider the evolution of $\chi(\rho_i^X)$ and $S(\rho_i)$. Initially, $\chi(\rho_0^X) = S(\rho_0) = 0$, since Bob begins in a state independent of X . Now, consider how $\chi(\rho_i^X)$ and $S(\rho_i)$ change for each of the four processes above.

1. This does not affect ρ_i^X and hence has no effect on $\chi(\rho_i^X)$ or $S(\rho_i)$.
2. It is easy to verify that χ and S are invariant under unitary transformations, so this does not affect $\chi(\rho_i^X)$ and $S(\rho_i)$ either.
3. Let B denote Bob's qubits after i steps and Q denote the qubit that Alice sends to Bob at the $i + 1^{\text{st}}$ step. By the subadditivity inequality and the fact that, for a single qubit Q , $S(Q) \leq 1$, $S(BQ) \leq S(B) + S(Q) \leq S(B) + 1$. Also, by the Araki-Lieb inequality [1], $S(BQ) \geq S(B) - S(Q) \geq S(B) - 1$. It follows that $S(\rho_{i+1}) \leq S(\rho_i) + 1$ and

$$\begin{aligned}
\chi(\rho_{i+1}^X) &= S(\rho_{i+1}) - \sum_{x \in \{0,1\}^n} p_x S(\rho_{i+1}^x) \\
&\leq (S(\rho_i) + 1) - \sum_{x \in \{0,1\}^n} p_x (S(\rho_i^x) - 1) \\
&= \chi(\rho_i^X) + 2.
\end{aligned} \tag{31}$$

4. In this case, ρ_{i+1}^X is ρ_i^X with one qubit traced out. It is known that tracing out a subsystem of any quantum system does not increase χ [15], so $\chi(\rho_{i+1}^X) \leq \chi(\rho_i^X)$. Note also that $S(\rho_{i+1}) \leq S(\rho_i) + 1$ for this process, by the Araki-Lieb inequality [1].

Now, since $\chi(\rho_i^X)$ can only increase when Alice sends a qubit to Bob and by at most 2, Eq. (28) follows. Also, since $S(\rho_i)$ can only increase when one party sends a qubit to the other and by at most 1, Eq. (29) follows. This completes the proof of Theorem 2.

Quantum Recurrent Networks for Simulating Stochastic Processes

Michail Zak¹ & Colin P. Williams²

Ultracomputing Group, Mail Stop 126-347, Jet Propulsion Laboratory,
California Institute of Technology, Pasadena, CA 91109-8099
colin@solstice.jpl.nasa.gov, zak@solstice.jpl.nasa.gov

Abstract. We introduce the concept of quantum recurrent networks by incorporating classical feedback loops into conventional quantum networks. We show that the dynamical evolution of such networks, which interleave quantum evolution with measurement and reset operations, exhibit novel dynamical properties finding application in pattern recognition, optimization and simulation. Moreover, decoherence in quantum recurrent networks is less problematic than in conventional quantum network architectures due to the modest phase coherence times needed for network operation.

Introduction

Large scale classical simulations of stochastic processes require vast quantities of random numbers. However, since the pioneering work of Church, Turing, Post and Gödel, it has been known that classical computers can only compute *functions*. In other words, the class of tasks that can be accomplished with a classical computer is exactly equivalent to the class of computable functions. However, as there is no *function* for computing a true random number, classical computers can only feign randomness. The purported calls to the “random number generator” often seen in modern programming languages are, in reality, calls to a *pseudo-random* number generator. A pseudo-random number generator is a deterministic function whose successive outputs pass many of the statistical tests of randomness. Unfortunately, the sequence of outputs can also harbor subtle correlations that are not immediately apparent from the common statistical measures of randomness.

To illustrate this point vividly, consider the RANDU “linear congruential generator”, a notoriously bad pseudo-random number generator that was common on IBM mainframes of the 1960s. A linear congruential generator is defined by:

$$N_{k+1} = (\ell N_k + m) \bmod n$$

¹ Supported by the JPL Information and Computing Technologies Research Section.

² Supported by the NASA/JPL Center for Integrated Space Microsystems and by the JPL Information and Computing Technologies Research Section.

C.P. Williams (Ed.): QCQC'98, LNCS 1509, pp. 75-88, 1999.

© Springer-Verlag Berlin Heidelberg 1999

where ℓ, m, n are fixed integers and $k = 1, 2, 3, \dots$. The resulting sequence of numbers, appears, superficially, to generate a set of random samples from a uniform distribution that lie in the range 0 to $n-1$ inclusive. We say “superficially” in the sense that, the sequence of numbers N_1, N_2, \dots passes many statistical tests of randomness. However, there is a subtle correlation lurking amongst these numbers that becomes apparent if you use them to choose a set of (supposedly) random points in a high dimensional space. In particular, if, as in RANDU, $\ell = 65539$, $n = 2^{31}$, $m = 0$ and $N_0 = 1$ then successive triples produced by the generator, N_k, N_{k+1}, N_{k+2} , can be taken to define the x -, y - and z -coordinates of a point in a 3-dimensional space. These points are plotted in fig.1 below from different viewing angles. From most viewing angles the points appear to be randomly distributed. But from a particular viewing angle you can see that they are not at all randomly distributed. In fact they lie in a set of parallel planes. Thus the sequence of supposedly “random” numbers output from the linear congruential generator are not random at all and could give misleading results if used in a numerical simulation of a stochastic process.

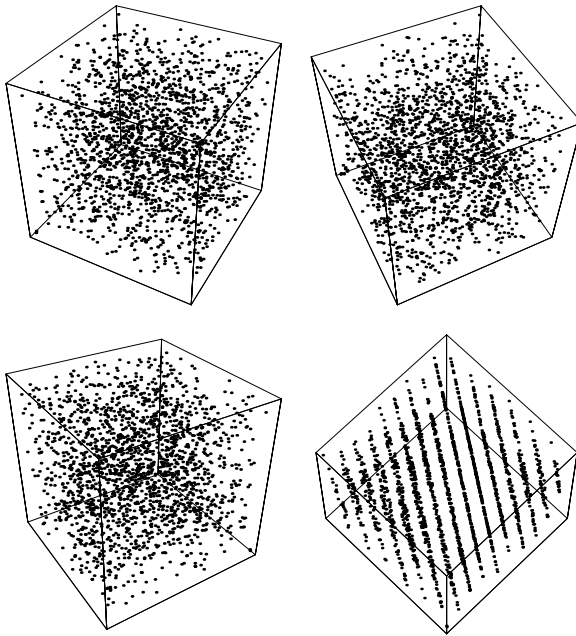


Fig.1 The same cube of points seen from four different viewpoints

Pseudo-random number generators are better today than they were in 1960. But one should not be complacent. As recently as 1992 bugs in a supposedly “good” pseudo-random number generator were discovered when a numerical simulation of an Ising spin system was performed as a test of the simulator against a known benchmark for which analytical results were known [Ferrenberg et al. 92].

One way to correct the problems inherent with pseudo-random number generators is to build a generator that exploits a truly random process itself. Strictly speaking, there is no such thing as randomness in classical physics. Nevertheless, certain classical dynamical systems exhibit a high degree of instability that would seem to make them reasonable candidates for effectively random number generators. Unfortunately, yet again one must be careful. Recently a study of the use of the logistic map, one of the simplest generators of deterministic chaos, has revealed that the chaotic sequence of numbers output does not possess exactly the same statistical properties as a truly random sequence[Phatak & Rao 95]. The deviation might be quite slight, but in situations where billions upon billions of random numbers are needed it could lead a Monte Carlo simulation astray.

The moral is that you cannot use classical physics to generate truly random numbers. However, in quantum physics, the non-deterministic outcome of a measurement made on a system in a superposed quantum state is, *as a matter of principle*, random. Unfortunately, quantum non-determinism is generally regarded as being of lesser importance than other quantum phenomena such as quantum interference and entanglement. This is partly because many people believe, mistakenly, that pseudo-random number generators are “good enough” and partly because the impressive speedups exhibited by quantum algorithms for factoring composite integers and for finding an item in an unsorted database, are due to interference and entanglement effects rather than non-determinism. However, such a dismissal of quantum non-determinism is premature. No matter how good new pseudo-random number generators are purported to be, their adequacy can only be assessed empirically within the context of a specific application. Moreover, the key quantum effect on which quantum cryptography depends is quantum non-determinism. We argue that as quantum non-determinism is, intrinsically, a random process, it provides a much better basis for the design of a random number generator.

It is easy to define a quantum procedure for selecting random integers in the range 0 to $2^n - 1$ by preparing n qubits in the state $|0\rangle|0\rangle|0\rangle\cdots|0\rangle$, applying the Walsh-Hadamard transform to each qubit separately, creating an equal superposition of all possible states of the register, and then reading the memory register. Once you have a mechanism for generating uniformly distributed random numbers you can create a generator for any other distribution using a function transformation[Tuckwell 95].

However true random number generation is not the same as true stochastic process generation. For example, in a Markov process the probability of obtaining a particular outcome for the next state depends upon the identity of the last state visited. By contrast the sequence of outcomes from a quantum random number generator are independent, identically distributed random variables. It is therefore interesting to ask whether there is a more *direct* way of using quantum mechanics to simulate stochastic processes?

Quantum Recurrent Networks

We can begin by asking what general features must such a simulator possess? First, we need to be able to generate a sequence of classically observable samples. This suggests that we are going to have to imagine a quantum device that allows repeated measurements. Second, we need to be able to bias the probability that a given state will appear as the next output given knowledge of some or all of the previous outputs. This is because, by definition, stochastic processes possess such a memory effect. The simplest way to accommodate such a memory effect is to imagine that the device is reset in a new state that somehow takes account of the states visited so far. These considerations lead naturally to our notion of a “quantum recurrent network”.

A quantum recurrent network consists of a conventional quantum network augmented with a classical measurement and quantum reset operation. The design of a one dimensional quantum recurrent network is shown in Fig.2.

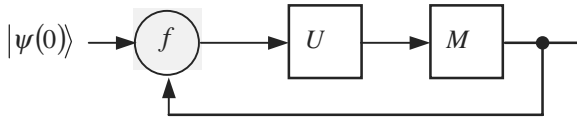


Fig. 2. A One dimensional quantum recurrent network

An initial state, $|\psi\rangle$, is fed into the network, transformed under the action of a unitary operator, U , subjected to a measurement, indicated by the measurement operator $M\{\}$, and the result of the measurement is used to control the new state fed back into the network at the next iteration. One is free to record, duplicate or even monitor the sequence of measurement outcomes, as they are all merely bits and hence constitute classical information. Moreover, one is free to choose the function used during the reset phase, including the possibility of adding no offset state whatsoever. Such flexibility makes the QRN architecture remarkably versatile. To simulate a Markov process, it is sufficient to return just the last output state to the next input at each iteration.

Quantum Reset Operation

The reset operation can be accomplished using conditional quantum logic. The basic strategy is to condition the operation performed on the offset state $|\psi\rangle$ upon the last measured outcome, $|i\rangle$. For example, $|\psi\rangle$ describes the state of 3 qubits, then the reset circuit will have the form shown in Fig. 3.

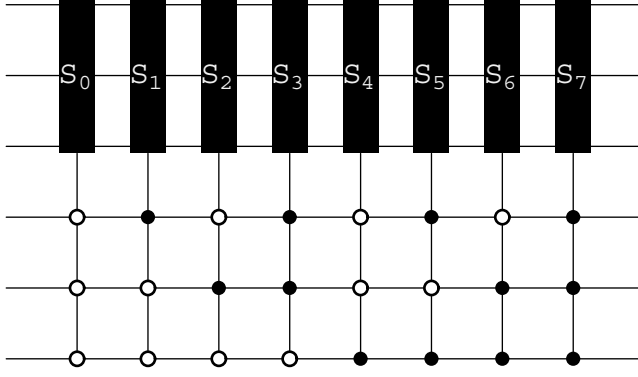


Fig. 3. Most general quantum reset circuit.

Here, a white dot represents the binary value 0 and a black dot the binary value 1. Mathematically, such a reset circuit is described by a 64×64 block diagonal unitary matrix in which each block, S_i , is an 8×8 unitary matrix.

$$\text{Reset Circuit} \equiv \begin{pmatrix} S_0 & & & & & & & \\ & S_1 & & & & & & \\ & & S_2 & & & & & \\ & & & S_3 & & & & \\ & & & & S_4 & & & \\ & & & & & S_5 & & \\ & & & & & & S_6 & \\ & & & & & & & S_7 \end{pmatrix}$$

Thus if the last measured output were, say, $|i\rangle = |011\rangle$ then the state re-entering the circuit at the next iteration will be $S_3|\psi\rangle$. An unfortunate drawback of the reset scheme is that the circuit contains an exponential number of gates. Later, we shall see that it is possible to invent alternative reset strategies that do not require exponentially sized circuits.

Network Dynamics

By design, we imbue the quantum recurrent network with a discrete time evolution according to the equation:

$$|\psi(t + \Delta t)\rangle = f\left(M\{U|\psi(t)\rangle\}, |\psi(0)\rangle\right)$$

where $|\psi(t)\rangle$ is the input to the network at time t , U is a unitary operator defined by $U = \exp(iH\Delta t/\hbar)$, and M is a measurement operator (in the computational basis) that has the effect of projecting the evolved state $U|\psi(t)\rangle$ into one of the eigenvectors of M . The curly brackets are intended to emphasize that M is to be taken as a measurement operation not a matrix product. In the simplest case, the combination function f , can be merely addition followed by renormalization, giving the specialized update rule:

$$|\psi(t + \Delta t)\rangle = \left\| M\{U|\psi(t)\} + |\psi(0)\rangle \right\|$$

where the notation $\|\cdot\|$ represents renormalization. As each iteration involves a measurement and reset operation, decoherence, the phenomenon that bedevils most hypothetical quantum computations, can be largely ignored, as the quantum recurrent networks need only operate coherently in between successive measurement operations; an interval of duration Δt .

In general, if one were to record the sequence measurement outcomes, it would not settle down to a predictable pattern. Instead, the sequence would hop about erratically between a finite, but possibly exponentially large, number of states, executing a true stochastic process. If the initial, i.e., offset, state vector is

$$|\psi(0)\rangle = \begin{pmatrix} a_0(0) \\ a_1(0) \\ \vdots \\ a_N(0) \end{pmatrix}$$

and M is a measurement operator in the computational basis, then $|\psi(t + \Delta t)\rangle$, the recurrent state re-entering the circuit, must consist of the sum of the offset state, $|\psi(0)\rangle$ plus quantum state (constructed afresh), $|i\rangle$, that corresponds to the last measured outcome, i . Hence, the recurrent state takes one of the forms:

$$\begin{aligned} |\phi_0\rangle &= \frac{1}{\sqrt{R_0}} \begin{pmatrix} 1 + a_0(0) \\ a_1(0) \\ \vdots \\ a_{N-1}(0) \end{pmatrix} = \frac{1}{\sqrt{R_0}} \begin{pmatrix} a_0^{(0)} \\ a_1^{(0)} \\ \vdots \\ a_{N-1}^{(0)} \end{pmatrix} \\ |\phi_1\rangle &= \frac{1}{\sqrt{R_1}} \begin{pmatrix} a_0(0) \\ 1 + a_1(0) \\ \vdots \\ a_{N-1}(0) \end{pmatrix} = \frac{1}{\sqrt{R_1}} \begin{pmatrix} a_0^{(1)} \\ a_1^{(1)} \\ \vdots \\ a_{N-1}^{(1)} \end{pmatrix} \\ &\vdots \end{aligned}$$

$$|\phi_{N-1}\rangle = \frac{1}{\sqrt{R_{N-1}}} \begin{pmatrix} a_0(0) \\ a_1(0) \\ \vdots \\ 1 + a_{N-1}(0) \end{pmatrix} = \frac{1}{\sqrt{R_{N-1}}} \begin{pmatrix} a_0^{(N-1)} \\ a_1^{(N-1)} \\ \vdots \\ a_{N-1}^{(N-1)} \end{pmatrix}$$

with re-normalization factors:

$$\begin{aligned} R_0 &= |1 + a_0(0)|^2 + |a_1(0)|^2 + \dots \\ R_1 &= |a_0(0)|^2 + |1 + a_1(0)|^2 + \dots \\ &\vdots \\ R_{N-1} &= |a_0(0)|^2 + |a_1(0)|^2 \dots + |1 + a_{N-1}(0)|^2 \end{aligned}$$

Thus, the recurrent (quantum) states entering the circuit and the measured (classical) outcomes follow the same Markov process. The transition probability matrix, T_1 , for this process is given by examining how each of the recurrent states, $|\phi_0\rangle \dots |\phi_{N-1}\rangle$ evolve under the action of U :

$$T_1 = \begin{pmatrix} \left| \frac{b_0^{(0)}}{\sqrt{R_0}} \right|^2 & \left| \frac{b_1^{(0)}}{\sqrt{R_0}} \right|^2 & \left| \frac{b_2^{(0)}}{\sqrt{R_0}} \right|^2 & \dots \\ \left| \frac{b_0^{(1)}}{\sqrt{R_1}} \right|^2 & \left| \frac{b_1^{(1)}}{\sqrt{R_1}} \right|^2 & \left| \frac{b_2^{(1)}}{\sqrt{R_1}} \right|^2 & \dots \\ \vdots & \vdots & \vdots & \ddots \\ \left| \frac{b_0^{(N-1)}}{\sqrt{R_{N-1}}} \right|^2 & \left| \frac{b_1^{(N-1)}}{\sqrt{R_{N-1}}} \right|^2 & \dots & \left| \frac{b_{N-1}^{(N-1)}}{\sqrt{R_{N-1}}} \right|^2 \end{pmatrix}$$

where

$$b_j^{(i)} = \sum_{\ell=0}^{N-1} U_{j\ell} a_\ell^{(i)} = U_{ji} + \sum_{\ell=0}^{N-1} U_{j\ell} a_\ell(0).$$

T_1 specifies, therefore, a classical transition probability matrix between a set of quantum states (the recurrent states) or equivalently between a set of classical states (the measurement outcomes).

Stochastic Attractors

The process defined by the transition probability matrix T_1 generates a *truly* random sequence of eigenstates. Although, for a given choice of U and $|\psi(0)\rangle$ every realization of the process will, in general, yield a different sequence of states, the statistical properties of these sequences, such as the frequencies of the various states visited, will eventually converge to a fixed distribution. This fixed distribution, is called a

stochastic attractor, and may be calculated as the fixed point of T_1 acting on any re-entrant state $|\phi_i\rangle$.

Thus the quantum recurrent network provides a mechanism for generating true stochastic attractors. Our model uses neither pseudo-random number generators nor classical white noise. The time taken to converge to this attractor is governed by the size of the largest eigenvalue of the transition matrix T_1 .

To be useful for Monte Carlo simulation, one would like to be able to tailor the quantum recurrent network so that it generates stochastic attractors that have prescribed characteristics. This can be accomplished by a process called “learning”. “Learning” consists of adapting model parameters until the quantum recurrent network produces the desired stochastic attractor to within an acceptable tolerance. This can be accomplished by varying the initial state fed into the network, $|\psi(0)\rangle$, the duration of the coherent evolution phase, Δt , or by selecting a different unitary matrix, U . From a practical perspective, as U will be embodied in physical hardware, it will be easier to perform learning by varying just $|\psi(0)\rangle$ and/or Δt . Varying Δt does, of course, change the unitary transformation applied during the coherent evolution phase, so it achieves a similar effect to picking a different U .

An arbitrary $N \times N$ dimensional unitary matrix has exactly N^2 free parameters. Therefore, in principle, a one dimensional QRN can generate a stochastic attractor having up to N^2 degrees of freedom. The method used to find a unitary matrix U that will generate a stochastic process with the desired properties, could be based on analytic minimization, gradient descent or genetic algorithms[Hertz, et al. 91].

Simulating an Eight State Markovian Process

Suppose we want to generate an 8 state Markov process that would generate the stochastic attractor shown in Fig. 4.

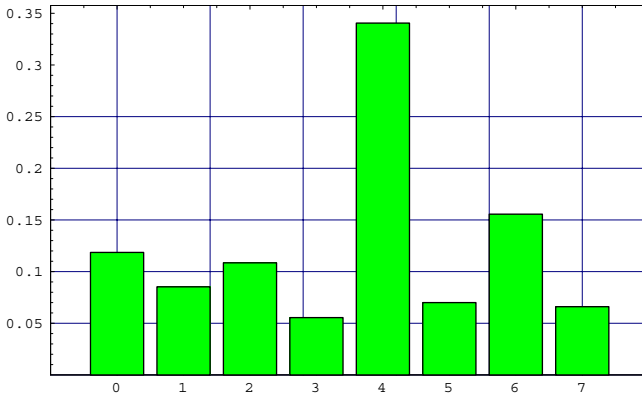


Fig. 4. Stochastic attractor defined on 8 states.

For this attractor, the probabilities of seeing states 0 through 7 are given by 0.119, 0.085, 0.108, 0.055, 0.341, 0.070, 0.156, and 0.066 respectively. As the attractor is one dimensional it can be obtained using a one dimensional QRN acting on 3 qubits.

To design such a QRN, we begin with a random unitary matrix, \hat{U}

$$\hat{U} = \begin{pmatrix} -0.322+0.197i & .027+.235i & -.283-.008i & .105-.389i & -.344-.32i & -.112-.385i & .145-.059i & .369+.128i \\ .105+.285i & .072+.342i & .182+.228i & .241+.216i & -.148-.209i & -.182+.352i & .465+.325i & -.169+.151i \\ .252+.206i & -.345-.155i & .31+.326i & .02+.37i & .088-.143i & .005-.433i & .116-.341i & .245+.077i \\ .018-.508i & .061+.162i & -.213-.032i & -.405+.251i & .167-.16i & -.308+.203i & .276-.184i & .336+.161i \\ -.141-.236i & -.189-.333i & -.147-.117i & .149+.297i & -.565+.353i & -.229-.15i & .19+.042i & -.197+.192i \\ -.08-.335i & -.253+.498i & .065+.278i & -.178-.103i & -.15+.164i & .337-.187i & .239-.083i & -.274-.33i \\ .221-.338i & .205+.041i & -.101+.625i & .177-.099i & -.123+.003i & .02-.032i & -.417+.111i & .024+.383i \\ .18+.105i & .172+.342i & -.136-.222i & .019+.417i & -.065+.345i & .314-.189i & -.103+.339i & .429-.007i \end{pmatrix}$$

Given \hat{U} , we can “train” the QRN to generate the desired stochastic attractor simply by varying the offset state, $|\psi(0)\rangle$. The training is complete when the difference between the QRN’s stochastic attractor and the target stochastic attractor is less than some threshold. For the given \hat{U} and the given target attractor, a satisfactory offset state vector is given by:

$$|\psi(0)\rangle = (-.093-.368i)|000\rangle - (.389+.311i)|001\rangle - (.029-.114i)|010\rangle + (.109-.221i)|011\rangle + (.325-.347i)|100\rangle - (.290+.162i)|101\rangle - (.330-.165i)|110\rangle + .244|111\rangle$$

1000 iterations of the QRN having this \hat{U} and $|\psi(0)\rangle$ induces the stochastic attractor shown in Fig. 5.

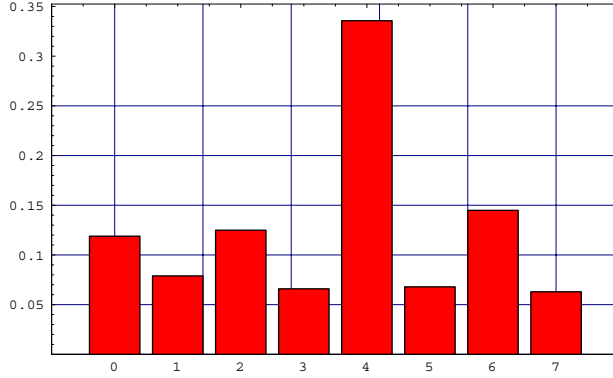


Fig. 5. The attractor generated by the “trained” QRN.

As you can see the agreement between the target attractor and the QRN’s attractor is very good. Moreover, one also learns a plausible transition probability matrix for the Markov chain that induces the target attractor.

$$T_1 = \begin{pmatrix} .243 & .009 & .236 & .088 & .121 & .045 & .201 & .057 \\ .173 & .003 & .090 & .069 & .151 & .343 & .048 & .123 \\ .147 & .025 & .319 & .061 & .168 & .066 & .155 & .059 \\ .026 & .033 & .243 & .187 & .371 & .022 & .066 & .052 \\ .129 & .152 & .028 & .001 & .537 & .036 & .095 & .022 \\ .108 & .036 & .010 & .203 & .244 & .069 & .118 & .021 \\ .016 & .139 & .030 & .021 & .319 & .028 & .409 & .038 \\ .059 & .052 & .148 & .063 & .378 & .082 & .048 & .170 \end{pmatrix}$$

This transition probability matrix is not unique because, if the attractor contains N degrees of freedom, the corresponding transition probability matrix contains $N^2 - N$ degrees of freedom. Thus the transition probability matrix is not as constrained as the attractor itself.

The k -Parallel Case

Next we generalize the concept of a quantum recurrent network to the case in which there are k networks working in parallel (see Fig. 6).

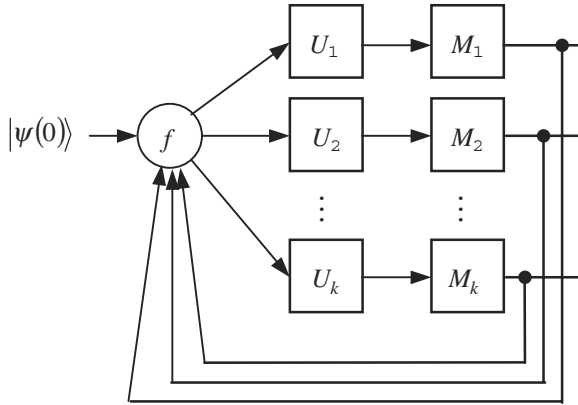


Fig. 6. The k -parallel quantum recurrent network

During the quantum evolution and measurement phases each network acts independently of the rest. However, during the reset operation the results of all the measurements are combined with the initial state to yield k identical input states. Note that the reset operation does not require an *arbitrary* quantum state to be cloned (which is impossible). Instead it only requires that k classical states, the outcomes of the k independent measurements, be copied. As this information is purely classical it can be copied freely. Moreover, the initial state $|\psi(0)\rangle$ is known and can be generated afresh as needed by each of the k networks. As a result, the feedback process we

propose is guaranteed to be physically realizable. The purpose of moving to the k -parallel quantum recurrent network is to permit us to generate *multi-dimensional* stochastic attractors. The reset operation gives us the flexibility to introduce correlations between the attractors in each dimension.

For the k -parallel QRN, the elements of the transition probability matrix now define the probability of making transitions between *sets* of measurement outcomes. The state entering each of the k networks at each iteration will have a form such as:

$$|\phi_{(i_1 i_2 \dots i_k)}\rangle = \frac{1}{\sqrt{R_{(i_1 i_2 \dots i_k)}}} \begin{pmatrix} a_0^{(i_1 i_2 \dots i_k)} \\ a_1^{(i_1 i_2 \dots i_k)} \\ \vdots \\ a_{N-1}^{(i_1 i_2 \dots i_k)} \end{pmatrix}$$

where the sequence $i_1 i_2 \dots i_k$ specifies the last ordered set of measurement outcomes obtained from the k networks and $R_{i_1 i_2 \dots i_k}$ is the renormalization constant given by:

$$R_{(i_1 i_2 \dots i_k)} = |a_0^{(i_1 i_2 \dots i_k)}|^2 + |a_1^{(i_1 i_2 \dots i_k)}|^2 + \dots + |a_{N-1}^{(i_1 i_2 \dots i_k)}|^2$$

The mathematical form of the amplitude $a_\ell^{(i_1 i_2 \dots i_{N-1})}$ depends upon how many of the components in the k -parallel network produced the same measurement outcome at the last iteration through the QRN i.e. how many of the i_ℓ in the sequence $i_1 i_2 \dots i_k$ were the same. If outcome i_ℓ is obtained n_{i_ℓ} times we have $a_\ell^{(i_1 i_2 \dots i_k)} = n_{i_\ell} + a_\ell(0)$.

As there are k networks and each network can produce one of N outcomes (independently), the k -parallel transition matrix defines a mapping from N^k distinct sets of input states to N^k sets of output states. If we denote the probability of transitioning from the set of inputs $i_1 i_2 \dots i_k$ to the set of outputs $j_1 j_2 \dots j_k$ by $p_{j_1 \dots j_k}^{(i_1 \dots i_k)}$ we have:

$$p_{j_1 \dots j_k}^{(i_1 \dots i_k)} = \left| \frac{U_1 b_{j_1}^{(i_1 i_2 \dots i_k)}}{\sqrt{R_{(i_1 i_2 \dots i_k)}}} \right|^2 \left| \frac{U_2 b_{j_2}^{(i_1 i_2 \dots i_k)}}{\sqrt{R_{(i_1 i_2 \dots i_k)}}} \right|^2 \dots \left| \frac{U_k b_{j_k}^{(i_1 i_2 \dots i_k)}}{\sqrt{R_{(i_1 i_2 \dots i_k)}}} \right|^2$$

where

$$U b_j^{(i_1 i_2 \dots i_k)} = \sum_{s=1}^k U_{j i_s} + \sum_{\ell=0}^{N-1} U_{j \ell} a_\ell(0)$$

Thus the k -parallel transition probability matrix has a tensor structure of the form $T_k = \{p_{j_1 j_2 \dots j_k}^{(i_1 i_2 \dots i_k)}\}_{N^k \times N^k}$ where the sequences $i_1 i_2 \dots i_k$ and $j_1 j_2 \dots j_k$ are defined with respect to some consistent ordering.

For the k -parallel architecture, there are $k N^2$ free parameters. Thus we ought to be able to generate k -dimensional stochastic attractors having up to $k N^2$ degrees of freedom.

An interesting corollary of the QRN dynamics concerns the dynamical behavior of two $k=1$ QRNs in comparison to a single $k=2$ QRN that combines them both. For simplicity we can set the initial (offset) state vector $|\psi(0)\rangle$ to be zero. Con-

sidered separately, the resulting stochastic processes have transition probabilities $p_{j_1}^{(i_1)}$ and $p_{j_2}^{(i_2)}$ given by:

$$p_{j_1}^{(i_1)} = \left| \frac{U_{j_1 i_1}^{(1)}}{\sqrt{2}} \right|^2, \quad p_{j_2}^{(i_2)} = \left| \frac{U_{j_2 i_2}^{(2)}}{\sqrt{2}} \right|^2$$

By contrast the transition probability matrix of the joint QRN has components:

$$p_{j_1 j_2}^{(i_1 i_2)} = \left| \frac{U_{j_1 i_1}^{(1)} + U_{j_1 i_2}^{(2)}}{\sqrt{2}} \right|^2 \times \left| \frac{U_{j_2 i_1}^{(1)} + U_{j_2 i_2}^{(2)}}{\sqrt{2}} \right|^2$$

Clearly $p_{j_1 j_2}^{(i_1 i_2)} \neq p_{j_1}^{(i_1)} \times p_{j_2}^{(i_2)}$ in general. Thus the two dimensional stochastic attractor generated by a 2-parallel quantum recurrent network is not simply the product of two 1-dimensional stochastic attractors.

Alternative Reset Strategies

So far, we have described the most general quantum recurrent network, i.e., one which involves an *arbitrary* offset state and an *arbitrary* unitary operator. Unfortunately, for such QRNs, the reset operation requires a circuit, like that depicted in Fig. 3, which is exponential in the number of qubits. This is because the required reset is different depending on which of the 2^n possible outcomes was obtained at the last measurement. Moreover, to implement an arbitrary unitary matrix as a quantum circuit could, in the worst case, require an *exponential* number of quantum gates [Knill 95]. Both these shortcomings can be sidestepped, however, by a slight modification to the QRN.

Instead of allowing *any* multi-particle state vector to serve as the offset vector we could allow only product states. This would enable the reset operation to be achieved in only a polynomial number of operations. For example, suppose the offset vector for a 2-qubit QRN is the product state $|\psi\rangle_1 |\phi\rangle_2$, then, instead of the reset operation with exponential cost, i.e., if a and b are the most recently measured classical outcomes, $|\psi\rangle_1 |\phi\rangle_2 \mapsto \left\| |\psi\rangle_1 |\phi\rangle_2 + |a\rangle_1 |b\rangle_2 \right\|$, we could impose a qubit-by-qubit reset operation $|\psi\rangle_1 |\phi\rangle_2 \mapsto \left\| (|\psi\rangle_1 + |a\rangle_1)(|\phi\rangle_2 + |b\rangle_2) \right\|$. Although the latter operation is conditional too, the conditioning is with respect to each individual qubit rather than the state of the entire multi-qubit register. Thus, the polynomial cost reset circuit would have the form shown in Fig. 7.

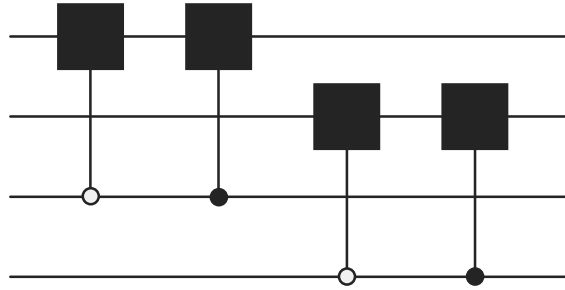


Fig. 7 A polynomially sized reset circuit.

Alternatively, we can imagine another reset strategy inspired by Brassard et al's discovery of a notion of the "closest" product state to an arbitrary entangled state [Brassard & Mor 98]. As we know the ideal entangled state that we should like to feed back into the QRN, we could instead use the closest product state to it. This would perhaps be the best compromise reset, as it would approximate the generic QRN and yet still only require a polynomial cost reset operation.

Likewise, rather than allowing U to be an arbitrary unitary operation, we could require that U be a unitary operation that is implementable in a polynomially sized quantum circuit. That is, we restrict consideration to unitary operators U that can be factored in the form $U = (U_1 \otimes U_2 \otimes \dots) \cdot (V_1 \otimes V_2 \otimes \dots) \dots$ where U_i and V_i are simple unitary operators describing 1-qubit or 2-qubit gates and the total number of terms in the product is bounded by a polynomial in the number of qubits. As a given stochastic attractor can, in general, be obtained from several different transition probability matrices, we have some degree of latitude over the exact choice of U . Thus restricting attention to U 's having a special (compact) decomposition ought not to be that limiting.

Summary

Quantum recurrent networks provide a mechanism for generating *true* stochastic attractors. By a process of learning we can tune the free parameters in a QRN to produce stochastic attractors having prescribed characteristics, such as a specific frequency distribution for the states visited. Moreover, as the QRNs operate by interleaving quantum evolution with measurement and reset operations, they are far less sensitive to decoherence than other designs of quantum computers.

Stochastic attractors find a wide range of applications in the physical and computational sciences. For example, one could use quantum recurrent networks as associative memories. Different stimuli, represented by different inputs $|\psi(0)\rangle$ would induce different stochastic attractors. The capacity of such quantum associative memories i.e. the number of distinct stimuli that can be recognized without unacceptable error, is much higher than for a comparable classical network.

Acknowledgements

The authors gratefully acknowledge financial support for this work from the NASA/JPL Center for Integrated Space Microsystems under contract 277-3R0U0-0 and from the JPL Information & Computing Technologies Research Section under contract 234-8AX24-0.

References

- [Brassard & Mor 98] G. Brassard & T. Mor, "Multi-particle Entanglement via 2-particle Entanglement", in Proceedings of the First NASA International Conference on Quantum Computing & Quantum Communications, NASA QCQC'98, Lecture Notes in Computer Science, Springer-Verlag (1998).
- [Ferrenberg & Landau 92] A. M. Ferrenberg & D. P. Landau, "Monte Carlo Simulations: Hidden errors from "Good" Random Number Generators", Physical Review Letters, Vol. 69, Number 23, 7th December (1992) pp3382-3384.
- [Hertz et al. 91] J. Hertz, A. Krogh & R. G. Palmer, "Introduction to the Theory of Neural Computation", Lecture Notes Volume I, Santa Fe Institute, Studies in Sciences of Complexity, Addison-Wesley Publishing Company (1991).
- [Knill 95] E. Knill, "Approximation by Quantum Circuits ", Los Alamos preprint, <http://xxx.lanl.gov/archive/quant-ph/9508006> (1995).
- [Phatak & Rao 95] S. C. Phatak & S. S. Rao, "Logistics Map: A Possible Random Number Generator", Physical review E, Volume 51, Number 4, April (1995) pp.3670-3678.
- [Tuckwell 95] H. Tuckwell, "Elementary Applications of Probability Theory", Second Edition, Chapman & Hall, Texts in Statistical Science (1995).
- [Williams & Clearwater 97] C. P. Williams & S. H. Clearwater, "Explorations in Quantum Computing", TELOS/Springer-Verlag, book/CD-ROM, ISBN 0-387-94768-X (1997)
- [Zak & Williams 98] M. Zak & C. P. Williams, "Quantum Neural Nets", International Journal of Theoretical Physics, Volume 37, Number 2, (1998) pp651-684.

Correlation between Correlations: Process and Time in Quantum Networks

Günter Mahler and Ilki Kim

Institut für Theoretische Physik, Universität Stuttgart Pfaffenwaldring 57
70550 Stuttgart, Germany

`mahler@theo.physik.uni-stuttgart.de`

Abstract. We study a special inhomogeneous quantum network consisting of a ring of M pseudo-spins (here $M = 4$) sequentially coupled to one and the same central spin under the influence of given pulse sequences (quantum gate operations). This architecture could be visualized as a quantum Turing machine with a cyclic “tape”. Rather than input-output-relations we investigate the resulting process, i.e. the correlation between one- and two-point expectation values (“correlations”) over various time-steps. The resulting spatiotemporal pattern exhibits many non-classical features including Zeno-effects, violation of temporal Bell-inequalities and quantum parallelism. Due to the strange web of correlations being built-up, specific measurement outcomes for the tape may refer to one or several preparation histories of the head. Specific families of correlation functions are more stable with respect to dissipation than the total wave-function.

1 Introduction

It has been shown that certain computational problems scale more favorably when carried out on a quantum system than on *any* classical computer (see e.g. [1]). The underlying “quantum complexity” may thus reduce computational complexity. Architectures for abstract quantum networks appropriate for such potential applications have been discussed (cf. e.g. [2]).

On the other hand, the control of such quantum networks appears to scale very badly with system size: In fact, in the language of statistical physics, that control would amount to use “micro-states” rather than “macro-states”, a challenging undertaking, indeed. It should therefore not come as a surprise that virtually all proposals up to now face severe problems when trying to go beyond the (coherent) control of something like $N = 10$ pseudo-spins [3,4,5,6]. Also the detailed theoretical simulation will become increasingly difficult if not eventually impossible beyond that limit. Fortunately, as will be shown below, even such small networks may show a surprisingly rich behavior in terms of correlation functions. Rather than the entanglement as such this pattern of correlations should be considered as the basis of the expected computational efficiency as well as other potential applications.

2 Composite Systems

2.1 States

The system we are going to investigate here is composed of $M + 1$ spins, $\mu = S, 1, 2, \dots, M$. The respective states are $|p(\mu)\rangle$, $p = 0, 1$. The corresponding product basis is $|u(M) \cdots r(2)q(1)p(S)\rangle \equiv |u \cdots rqp\rangle$. Arranged in the order of increasing binary numbers we also introduce the single-index notation $|s\rangle$, $s = 0, 1, \dots, 2^{M+1} - 1$ by identifying $|0\rangle = |0 \cdots 000\rangle$, $|1\rangle = |0 \cdots 001\rangle$, $|2\rangle = |0 \cdots 010\rangle$, etc. This single-index representation will not only serve as a means to simplify some algebra. It reminds us that one can entirely avoid talking about entanglement while, nevertheless, keeping the product-space background still operative, though in a more subtle way: In terms of the *specific* operator combinations and their expectation values.

2.2 Cluster-Operators

For $M + 1 = 5$ there are $(2^5)^2 = 1024$ orthogonal basis operators. One possible choice would be products of local transition-operators, $\hat{P}_{pq}(\mu) = |p(\mu)\rangle\langle q(\mu)|$. For reasons that will become clear shortly it is more convenient to separate out the local unit operators $\hat{1}(\mu)$ so that the remaining operators become traceless. Such a scheme is provided by the Hermitian and unitary $SU(2)$ -generators, $\hat{\lambda}_j(\mu)$,

$$\begin{aligned}\hat{\lambda}_1(\mu) &= \hat{P}_{01}(\mu) + \hat{P}_{10}(\mu) \\ \hat{\lambda}_2(\mu) &= i\hat{P}_{01}(\mu) - i\hat{P}_{10}(\mu) \\ \hat{\lambda}_3(\mu) &= \hat{P}_{11}(\mu) - \hat{P}_{00}(\mu) \\ \hat{\lambda}_0(\mu) &= \hat{P}_{11}(\mu) + \hat{P}_{00}(\mu) = \hat{1}(\mu).\end{aligned}\tag{1}$$

The corresponding product operators $(j, k, l, m, n = 0, 1, 2, 3)$ [\[7\]](#)

$$\hat{Q}_{jklmn} = \hat{\lambda}_j(S)\hat{\lambda}_k(1)\hat{\lambda}_l(2)\hat{\lambda}_m(3)\hat{\lambda}_n(4)\tag{2}$$

with $(\hat{Q}_{jklmn})^2 = \hat{1}$ for any (j, k, l, m, n) and

$$\text{Tr}\{\hat{Q}_{jklmn}\hat{Q}_{j'k'l'm'n'}\} = 2^5\delta_{jj'}\delta_{kk'}\delta_{ll'}\delta_{mm'}\delta_{nn'}\tag{3}$$

then come in 6 classes, depending on the number $c = 0, 1, \dots, M + 1$ of subsystems they act on, i.e. the number of indices unequal zero. $\hat{Q}_{00000} = \hat{1}$ is the only $c = 0$ cluster operator. When transcribed to the single index-space, $s = 0, 1, \dots, 2^5 - 1$, these operators appear like a set of “generalized” $SU(2)$ -operators of the form given in eq. [\(II\)](#) with each single transition or projection operator replaced by a group of $2^M = 16$. Such operator combinations would be hard if not impossible to implement in a simple one-particle system with 2^5 states; they reflect the structure of the underlying product space. Correspondingly, the expectation-value of any cluster-operator is a sum of 2^{M+1} density

matrix elements in the single-index space. Examples for $M = 2$ are the $c = 1$ -cluster operators,

$$\begin{aligned}\hat{Q}_{300} &= (\hat{P}_{11} + \hat{P}_{33} + \hat{P}_{55} + \hat{P}_{77}) - (\hat{P}_{00} + \hat{P}_{22} + \hat{P}_{44} + \hat{P}_{66}) \\ \hat{Q}_{030} &= (\hat{P}_{22} + \hat{P}_{33} + \hat{P}_{66} + \hat{P}_{77}) - (\hat{P}_{00} + \hat{P}_{11} + \hat{P}_{44} + \hat{P}_{55})\end{aligned}\quad (4)$$

or $c = 2$ -cluster operators such as,

$$\begin{aligned}\hat{Q}_{330} &= (\hat{P}_{00} + \hat{P}_{33} + \hat{P}_{44} + \hat{P}_{77}) - (\hat{P}_{11} + \hat{P}_{22} + \hat{P}_{55} + \hat{P}_{66}) \\ \hat{Q}_{303} &= (\hat{P}_{00} + \hat{P}_{22} + \hat{P}_{55} + \hat{P}_{77}) - (\hat{P}_{11} + \hat{P}_{33} + \hat{P}_{44} + \hat{P}_{66}) .\end{aligned}\quad (5)$$

Any operator \hat{A} in the 2^5 -dimensional Hilbert-space of spin-states can be represented as (summation over repeated indices),

$$\hat{A} = \frac{1}{2^5} A_{jklmn} \hat{Q}_{jklmn} \quad (6)$$

with the parameters

$$A_{jklmn} = \text{Tr}\{\hat{A}\hat{Q}_{jklmn}\} . \quad (7)$$

(Tr means trace over the total Hilbert-space.) In particular, the network Hamiltonian \hat{H} can be specified by the model parameters H_{jklmn} ; they are usually constrained to $c=0, 1$ and 2-cluster-terms [7]. The density operator $\hat{\rho}$ is uniquely defined by the set of expectation-values (note that \hat{Q}_{jklmn} is unitary)

$$-1 \leq K_{jklmn} = \text{Tr}\{\hat{\rho}\hat{Q}_{jklmn}\} \leq 1 \quad (8)$$

with c -cluster-operators defining c -particle correlations. For a pure state, $\hat{\rho} = |\psi\rangle\langle\psi|$, eq. (8) reduces to

$$K_{jklmn} = \langle\psi|\hat{Q}_{jklmn}|\psi\rangle . \quad (9)$$

By definition, $K_{00000} = 1$; the local Bloch-vectors $K_{j0000}, K_{0k000}, K_{00l00}$, etc. ($j, k, l = 1, 2, 3$) are equivalent to the respective reduced density matrices. A pure *local* state has Bloch-vector-length 1. For so-called product states all these correlations factor into one-point functions, i.e.

$$K_{jklmn} = K_{j0000} \cdot K_{0k000} \cdot K_{00l00} \cdot K_{000m0} \cdot K_{0000n}$$

but, in general, they are independent. Local realism (cf. [8]), to be sure, postulates that an appropriate distribution of local variables (eigenvalues $\lambda_j = \pm 1$) could explain *all* these correlation functions rendering them statistically dependent; at least for larger networks this approach is no longer tenable. On the other hand, as will be shown below, the quantum mechanical evolution generates “correlations between correlations”.

For later reference we also define symmetrized correlation functions *within* one and the same system μ :

$$C_{AB}^{(\mu)} = \frac{1}{2}(\text{Tr}\{\hat{\rho}\hat{A}(\mu)\hat{B}(\mu)\} + \text{Tr}\{\hat{\rho}\hat{B}(\mu)\hat{A}(\mu)\}) . \quad (10)$$

Restricting ourselves to traceless operators, this correlation is independent of $\hat{\rho}$ (for two-dimensional Hilbert-spaces) and can simply be written as the normalized scalar product between the two representing vectors [7]; for $\mu = S$, e.g.,

$$C_{AB}^{(S)} = \frac{1}{2^{10}} A_{j0000} B_{j0000} . \quad (11)$$

2.3 Unitary Transformations

A unitary transformation of an operator \hat{A} ,

$$\hat{A}' = \hat{U} \hat{A} \hat{U}^+ \quad (12)$$

with $\hat{U}^+ \hat{U} = \hat{U} \hat{U}^+ = \hat{1}$, reads in terms of the $SU(2)$ - parameters,

$$A'_{jklmn} = X_{j \ k \ l \ m \ n}^{j' \ k' \ l' \ m' \ n'} A_{j' k' l' m' n'} \quad (13)$$

where

$$X_{j \ k \ l \ m \ n}^{j' \ k' \ l' \ m' \ n'} = \frac{1}{2^5} \text{Tr} \{ \hat{U}^+ \hat{Q}_{jklmn} \hat{U} \hat{Q}_{j'k'l'm'n'} \} \quad (14)$$

(For $\hat{U} = \hat{1}$, X is just the unit matrix, see eq. (3)). There are different types: We may distinguish transformations which operate in certain subspaces only. The locally selective transformation $\hat{U}(S)$ in the 2-dimensional local Hilbert-space of S , e.g., is equivalent to a local rotation of the $SU(2)$ -parameters with respect to the first index j , generated by (cf. [9])

$$\begin{aligned} X_{j \ k \ l \ m \ n}^{j' \ k' \ l' \ m' \ n'} &= X_{jj'}^{(S)} \delta_{kk'} \delta_{ll'} \delta_{mm'} \delta_{nn'} \\ X_{jj'}^{(S)} &= \frac{1}{2} \text{Tr}_S \{ \hat{U}^+(S) \hat{\lambda}_j(S) \hat{U}(S) \hat{\lambda}_{j'}(S) \} . \end{aligned} \quad (15)$$

(Here, Tr_S means trace over the subspace of S only.) As $X_{00}^{(S)} = 1$, $X_{jj'}^{(S)} = 0$ if either j or j' is zero, all parameters A_{0klmn} are invariants. Correspondingly, a unitary transformation $\hat{U}(S, 1)$ leaves the expectation values A_{00lmn} unchanged, etc. These invariants (conservation laws) are important characteristics of the respective transformations.

2.4 Time

As we do not consider equations of motion explicitly, time enters at most indirectly: To specify change, order and duration. For closed systems, unitary transformations are the only allowed type of changes (of states or observables) in closed quantum systems. Typically they are generated by the underlying Hamilton model. In the Schrödinger-picture this unitary transformation is applied to $\hat{\rho}$, in the Heisenberg-picture the inverse transformation (replacing \hat{U} by \hat{U}^+ and vice versa) is applied to the observables.

Parameter time T will come in with respect to the *order*, in which certain transformations are applied, as a *continuous parameter* controlling the individual transformation *quantitatively* (“pulse length” t), and, eventually, with respect to the *order* of measurements. Finally, the induced dynamics can be characterized by correlation-and recurrence-times.

3 The Turing Model

Our system is sketched in Fig 3. Spin S is the “Turing head”, the other $\mu = 1, 2, \dots, M = 4$ subsystems denote memories (as part of a circular “Turing tape”); the latter do not interact directly and are separated by “empty” cells. The head interacts with at most one cell at a time [10]; it moves clockwise and step by step to one of the $2M$ positions on the tape; there is no need for a feed-back between the internal quantum state of the network and this pre-determined “classical” movement.

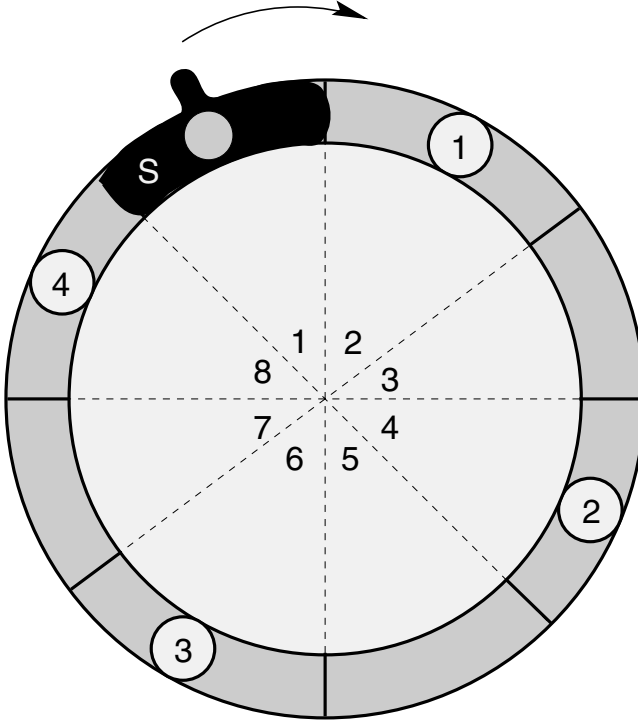


Fig. 1. Quantum Turing machine ($M = 4$). The circular Turing tape consists of $\mu = 1, 2, \dots, M$ memory cells (position 2μ) separated by empty cells (position $2\mu - 1$). The Turing head moves clockwise thus initiating a local (position-index odd) or a pair transformation, respectively (position index even).

We assume to have explicit control over the model parameters H_{jklmn} defining the Hamiltonian, which may even be modified in terms of pulses in parameter-time (t_j is the pulse-length):

$$\hat{H}(t) = \hat{H}_j \text{ for } T_{j-1} \leq t < T_{j-1} + t_j \equiv T_j. \quad (16)$$

Granted this access we can implement virtually any unitary transformation via

$$\hat{U}(T_{j-1} + t_j, T_{j-1}) \equiv \hat{U}_j = e^{-i\hat{H}t_j/\hbar} \quad (17)$$

(j is the step number), though this may seriously be limited in practice. If the Turing head is over an empty cell (tape position $2\mu - 1$), a local transformation $\hat{U}_\alpha(S)$ on S is applied, if it is in contact with a memory cell μ at position 2μ a pair transformation on (S, μ) is induced ($\mu = 1, 2, \dots, M$).

3.1 Local Transformation on (S)

Let us consider the one-parameter-form

$$\begin{aligned} |0(S) > &\longrightarrow \cos(\alpha/2)|0(S) > -i \sin(\alpha/2)|1(S) > \\ |1(S) > &\longrightarrow -i \sin(\alpha/2)|0(S) > + \cos(\alpha/2)|1(S) > \end{aligned} \quad (18)$$

which can be generated by ($M = 4$)

$$\hat{U}_\alpha(S) = \hat{Q}_{00000} \cos(\alpha/2) - \hat{Q}_{10000} i \sin(\alpha/2) = \hat{U}_{-\alpha}^+(S) . \quad (19)$$

According to eqs. (19) (15) and (2), we find

$$\begin{aligned} X_{jj'}^{(S)} &= \cos^2(\alpha/2) \delta_{jj'} + \frac{1}{2} \sin^2(\alpha/2) \text{Tr}_S \{ \hat{\lambda}_1 \hat{\lambda}_j \hat{\lambda}_1 \hat{\lambda}_{j'} \} \\ &\quad + \frac{i}{4} \sin \alpha \text{Tr}_S \{ \hat{\lambda}_1 \hat{\lambda}_j \hat{\lambda}_{j'} - \hat{\lambda}_j \hat{\lambda}_1 \hat{\lambda}_{j'} \} \end{aligned} \quad (20)$$

so that $X_{00}^{(S)} = X_{11}^{(S)} = 1$, $X_{22}^{(S)} = X_{33}^{(S)} = \cos \alpha$, $X_{32}^{(S)} = -X_{23}^{(S)} = \sin \alpha$. (Here and in the following all terms not explicitly given are zero.) This matrix $X_{ij}^{(S)}$ defines a rotation of the Bloch-vector of S around the $k=1$ -axis in the 2,3-plane. The phase α may be taken to result from a pulse of duration t

$$\alpha = gt \quad (21)$$

where g would be the coupling strength to an external optical driving field. The correlation function between $\hat{A} = \hat{\lambda}_3(S)$ transformed by ϕ and the same operator transformed by phase angle $\phi + \alpha$ then is, according to eq. (11),

$$C_{33}^{(S)}(\phi, \phi + \alpha) = \cos \alpha . \quad (22)$$

Based on eq. (21) this expectation value can be interpreted as a 2-time 1-particle correlation function in the Heisenberg-picture. Combinations of these have been shown to violate temporal Bell inequalities [11].

3.2 Pair Transformation on (S, μ)

This unitary transformation is taken as the conditioned π -pulse, ($q = 0, 1$)

$$\begin{aligned} \text{Resonance: } |0(S)0(\mu) > &\longleftrightarrow |0(S)1(\mu) > \\ \text{Off-resonance: } |1(S)q(\mu) > &\longleftrightarrow |1(S)q(\mu) > \end{aligned} \quad (23)$$

which we may write, in terms of cluster operators, in the form

$$\begin{aligned} \hat{U}(S, 1) &= \hat{P}_{00}(S)\hat{\lambda}_1(1) + \hat{P}_{11}(S)\hat{1}(1) \\ &= \frac{1}{2}(\hat{Q}_{00000} + \hat{Q}_{30000} + \hat{Q}_{01000} - \hat{Q}_{31000}) = \hat{U}^+(S, 1) . \end{aligned} \quad (24)$$

These operators $\hat{U}(S, \mu)$ commute; their implementation requires pair interactions, which make the transition frequency in subsystem μ depend on the state of subsystem S [7][12]. This transformation has become known as the (quantum-) controlled NOT [2], as subsystem S acts as a control for a π -pulse on μ . We may associate a fixed pulse duration t_0 with this implementation; here we assume $t_0 \approx 0$. In general, the two types of unitary operators do not commute:

$$[\hat{U}(S, \mu), \hat{U}_\alpha(S)] = \sin(\alpha/2) (\hat{1}(\mu) - \hat{\lambda}_1(\mu))\hat{\lambda}_2(S) . \quad (25)$$

4 The Process

4.1 The First Cycle

We are now in a position to follow up the ordered sequence of $2M = 8$ unitary transformations,

$$|\psi^{(1,j)} > = \hat{U}_j |\psi^{(1,j-1)} > \quad (26)$$

where $(\mu = 1, 2, \dots, M)$

$$\begin{aligned} \hat{U}_{2\mu-1} &= \hat{U}_{\alpha_\mu}(S) \\ \hat{U}_{2\mu} &= \hat{U}(S, \mu) . \end{aligned} \quad (27)$$

Here and in the following the upper index pair in parenthesis denotes the cycle number m and the step number j , respectively. With $j = 2\mu$ ($\mu = 1, 2, \dots, M$) we may associate the time (cf. eq. (21))

$$T_{2\mu} = \sum_{i=1}^{\mu} t_{2i-1} = \sum_{i=1}^{\mu} \alpha_i/g \approx T_{2\mu-1} . \quad (28)$$

T_{2M} is then the time needed for each cycle. Now, let the initial state be $|\psi^{(1,0)} > = |0 > = |00000 >$ so that the local Bloch-vectors are given by

$$K_{30000}^{(1,0)} = K_{03000}^{(1,0)} = \dots = K_{00003}^{(1,0)} = -1 . \quad (29)$$

In the first step we apply the local transformation with a phase α_1 leading to

$$|\psi^{(1,1)}\rangle = \cos(\alpha_1/2) |0\rangle - i \sin(\alpha_1/2) |1\rangle. \quad (30)$$

In the second step we execute the pair transformation on $(S, 1)$:

$$|\psi^{(1,2)}\rangle = \cos(\alpha_1/2) |2\rangle - i \sin(\alpha_1/2) |1\rangle. \quad (31)$$

In the third step we again apply the local transformation, now with phase α_2 , leading to

$$\begin{aligned} |\psi^{(1,3)}\rangle &= \cos(\alpha_1/2) \cos(\alpha_2/2) |2\rangle - i \cos(\alpha_1/2) \sin(\alpha_2/2) |3\rangle \\ &\quad - i \sin(\alpha_1/2) \cos(\alpha_2/2) |1\rangle - \sin(\alpha_1/2) \sin(\alpha_2/2) |0\rangle. \end{aligned} \quad (32)$$

In the “Heisenberg-picture”, this implies between step 2 and step 3 the local correlation as given by eq. (22) with $\alpha = \alpha_2$. In the 4th step the pair transformation on $(S, 2)$ implies

$$\begin{aligned} |\psi^{(1,4)}\rangle &= \cos(\alpha_1/2) \cos(\alpha_2/2) |6\rangle - i \cos(\alpha_1/2) \sin(\alpha_2/2) |3\rangle \\ &\quad - i \sin(\alpha_1/2) \cos(\alpha_2/2) |1\rangle - \sin(\alpha_1/2) \sin(\alpha_2/2) |4\rangle. \end{aligned} \quad (33)$$

This procedure is continued with respect to the next memory cells 3 and 4 (steps 5 through 8). We note that the single-subsystem expectation values of subsystem S and μ obey the relations

$$\begin{aligned} K_{30000}^{(1,2)} &= -K_{03000}^{(1,2)} = K_{30000}^{(1,0)} \cos \alpha_1 \\ K_{30000}^{(1,4)} &= -K_{03000}^{(1,4)} = K_{30000}^{(1,2)} \cos \alpha_2 \quad \text{etc.} \\ K_{10000}^{(1,2\mu)} &= K_{20000}^{(1,2\mu)} = 0 \end{aligned} \quad (34)$$

and as a consequence of the controlled-NOT-logic (cf. eq. (5)),

$$K_{33000}^{(1,2)} = K_{30300}^{(1,4)} = K_{30030}^{(1,6)} = K_{30003}^{(1,8)} = -1. \quad (35)$$

We thus see that the two systems, S and μ , are strictly anti-correlated after step 2μ (the state $|\psi^{(1,2)}\rangle$, e.g., is actually an eigenstate of \hat{Q}_{33000} !), while the local Bloch-vector-lengths are less than 1, i.e. local properties are not dispersion-free (“fuzzy”). This is typical for non-classical correlations. There can be *strict* correlations between *fuzzy* subsystems.

4.2 Cycles $m \geq 1$

We can summarize and generalize the above results by introducing the following functions:

$$\begin{aligned} \kappa^{(m, 2M)}(\alpha_1, \alpha_2, \dots, \alpha_j) &= \\ \frac{1}{2} [\cos(m\alpha_1) \cos(m\alpha_2) \cdots \cos(m\alpha_j)] &+ \frac{1}{2} \begin{cases} 1 & m \text{ even} \\ \cos \alpha_1 \cos \alpha_2 \cdots \cos \alpha_j & m \text{ odd,} \end{cases} \end{aligned}$$

$\kappa_s^{(m,2M)}$ as above with $\cos m\alpha_1$ replaced by $\sin m\alpha_1$, $\cos \alpha_1$ replaced by $-\sin \alpha_1$ and the 1 replaced by 0 ($j \leq M$),

$$\begin{aligned}\phi_k^{(m,2M)} &= -\cos(m\alpha_1/2) \cos(m\alpha_2/2) \cdots \cos(m\alpha_M/2) & m \text{ even,} \\ \phi_k^{(m,2M)} &= \cos((m+1)\alpha_1/2) \cos((m+1)\alpha_2/2) \cdots \cos((m+1)\alpha_k/2) \\ &\quad \times \cos((m-1)\alpha_{k+1}/2) \cdots \cos((m-1)\alpha_M/2) & m \text{ odd,}\end{aligned}$$

and $\chi_k^{(m,8)} = -\phi_k^{(m,8)}$ with $(m+1)$ replaced by $(m-1)$ and vice versa.

Then, at the end of each cycle m , the Turing head can be described by ($M=4$)

$$\begin{aligned}K_{10000}^{(m,j)} &= 0 \\ K_{20000}^{(m,8)} &= \kappa_s^{(m,8)}(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \\ K_{30000}^{(m,8)} &= -\kappa^{(m,8)}(\alpha_1, \alpha_2, \alpha_3, \alpha_4)\end{aligned} \quad (36)$$

and the memory cells by

$$\begin{aligned}K_{03000}^{(m,8)} &= \phi_1^{(m,8)} \\ K_{00300}^{(m,8)} &= \phi_2^{(m,8)} \quad \text{etc.} \\ K_{03300}^{(m,8)} &= \kappa^{(m,8)}(\alpha_2) \\ K_{00330}^{(m,8)} &= \kappa^{(m,8)}(\alpha_3) \\ K_{00033}^{(m,8)} &= \kappa^{(m,8)}(\alpha_4) \\ K_{03030}^{(m,8)} &= \kappa^{(m,8)}(\alpha_2, \alpha_3) \\ K_{00303}^{(m,8)} &= \kappa^{(m,8)}(\alpha_3, \alpha_4) \\ K_{03003}^{(m,8)} &= \kappa^{(m,8)}(\alpha_2, \alpha_3, \alpha_4) .\end{aligned} \quad (37)$$

The memory pair-correlations are all positive for m even and decay with “step distance”, i.e. the number of intermediate rotation and coupling steps to other memory cells (cf. also Fig 4.2). The pair correlations between Turing head and the memories are given by

$$\begin{aligned}K_{33000}^{(m,8)} &= \chi_1^{(m,8)} \\ K_{30300}^{(m,8)} &= \chi_2^{(m,8)} \quad \text{etc.}\end{aligned} \quad (38)$$

All the expectation values are strictly periodic in m if $\alpha_j = 2\pi/p_j$ for all $j = 1, 2, \dots, M$ with p_j a whole number. The period p is then the smallest even number that has all these p_j as factors.

In a similar way one obtains the results for step numbers smaller than $2M = 8$. Generalizations to the situation where the phase angles differ from cycle to cycle are also straight-forward. For example, based on eqs. (37), (38) we find a web of correlations like

$$K_{33000}^{(m,j)} \cdot K_{03000}^{(m,j)} = K_{30300}^{(m,j)} \cdot K_{00300}^{(m,j)} = K_{30030}^{(m,j)} \cdot K_{00030}^{(m,j)} \quad \text{etc.} \quad (39)$$

valid for all steps j within any cycle m .

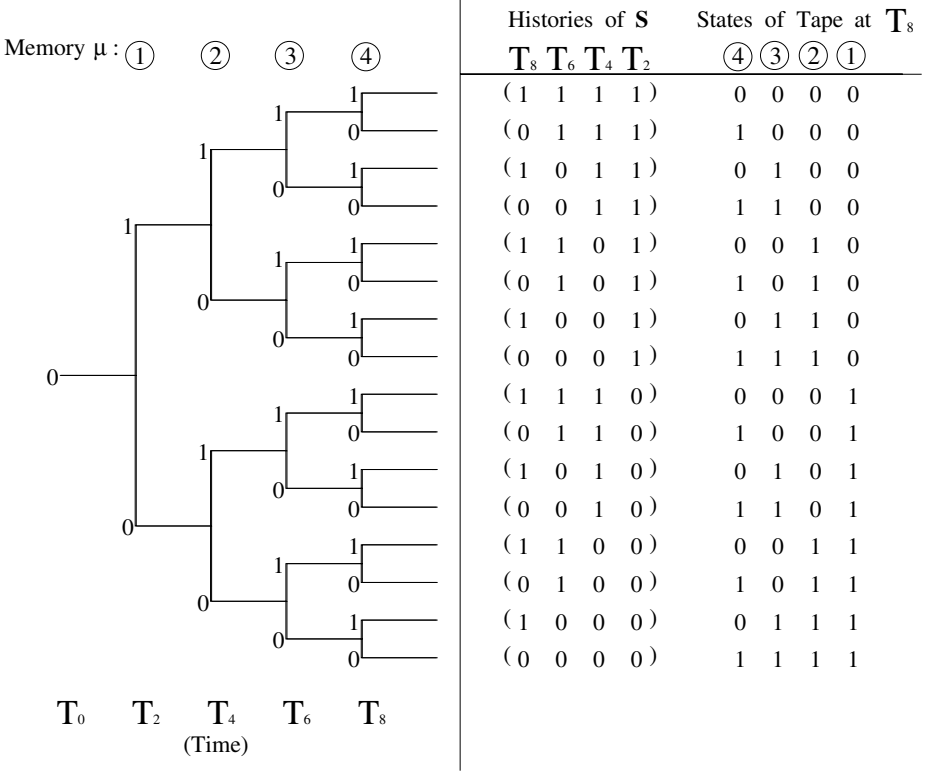


Fig. 2. Alternative histories. a. Decision tree with respect to step number $2\mu = 2, 4, 6, 8$, as realized in an *ensemble* of non-interacting spins S under the series of local transformations $\hat{U}_{\alpha\mu}(S)$, but with immediate *actual* measurements (replacing $\hat{U}(S, \mu)$ of our Turing machine) at times $T_{2\mu}$. b. For the single Turing machine all the possible histories are yet undecided and associated with the states of the Turing tape as given.

5 Reduced Descriptions

5.1 Turing-Head S

The description reduced to the subsystem S is based on the local Bloch-vector $K_{j0000}, j = 1, 2, 3$ only. Starting from the ground-state, $K_{30000}^{(1,0)} = -1$, this vector is subject to the rotation as given by eq. (19). We see that each controlled NOT operation implies a projection on the 3-axis ($K_{10000} = K_{20000} = 0$). The result of eq. (36) for cycle 1 is easily generalized to $M > 4$ with $\nu = 1, 2, \dots, M$ and $\alpha_\nu = \pi/M$. We find

$$K_{30000\dots}^{(1,2M)} = -\cos^M(\pi/M). \quad (40)$$

With $\alpha_\nu = gt_{2\nu-1}$ (cf. eq. (21); $2\nu - 1$ is the step number), the quantum-Zeno-effect [13],[14],[15] results within the fixed time $gT_{2M} = \pi$ (cf. eq. (28)). It is interesting to note that the reduced density matrix (or Bloch-vector) of subsystem S is, at any time t , identical with the density matrix of an *ensemble* of non-interacting spins (all with the same initial state and subject to the same local unitary transformation) but *actually measured* at each time $T_{2\nu}$, $\nu = 1, 2, \dots, M$. For each ensemble member the series of measurements constitutes a “decision-tree”, with each measurement result given by $K'_{30000} = \pm 1$ (see Fig 2). The ensemble average over these trajectories leads back to the behavior realized here by just one single object! The respective density matrices are identical. This is what one may call *quantum parallelism*. The interaction with the tape generates a dynamical evolution of the Turing head S equivalent to 2^M different histories (cf. [16]), clearly an exponential gain. This will only hold, though, as long as no measurements are performed.

5.2 Turing-Tape

Contrary to the Turing head S , the other subsystems are each addressed by unitary transformations only once (within each cycle). Due to the built-in logic the state of subsystem 1 is strictly anti-correlated with S after preparation step 2, subsystem 2 is anti-correlated with S after step 4, and so on. This means that an actual projective measurement performed on these subsystems would reveal also the respective states of S . When the transformations are interpreted to happen in parameter-time $T_{2\mu}$, the subsystems $\mu \neq S$ indeed act as a kind of “memory”. They allow delayed measurements on S . One may argue that this fact is the origin of the quantum-Zeno-effect discussed in Sect. (5.1): It suffices to *be able to measure* in order to get the freezing-tendency of measurements (“virtual watchdog”- effect).

Local measurements of the memory cells amounts to the application of a projection or transition-operator like $\hat{P}_{01}(\mu)$. As these operators commute among each other (for different μ) and with any of the unitary operators *not* acting on μ , we can postpone these measurements up to one cycle. For $\mu = 1$, e.g.,

$$\begin{aligned} \hat{U}(S, 4)\hat{U}_{\alpha_4}(S) \cdots \hat{P}_{01}(1)\hat{U}(S, 1)\hat{U}_{\alpha_1}(S)|\psi^{(m,0)} > = \\ \hat{P}_{01}(1)\hat{U}(S, 4)\hat{U}_{\alpha_4}(S) \cdots \hat{U}(S, 1)\hat{U}(S)_{\alpha_1}|\psi^{(m,0)} > . \end{aligned} \quad (41)$$

Let us first restrict ourselves to cycle $m = 1$ with its decision tree (Fig 4.2). The time order of these measurements (i.e. the measurement process) need not correspond to the time-order, in which the memory cells have been visited by the Turing head: The actual history for the latter (out of the possibilities as shown in Fig 4.2) may thus be “realized” even backward in time!

But not only this: The correlation between memory cell 1 and 2, e.g., must, by construction (cf. eq. (35)) and the invariance property $K_{03300}^{(1,8)} = K_{03300}^{(1,4)}$, reflect the correlation between the states of S taken at T_2 and T_4 , respectively. This is readily verified by comparing our result for $K_{03300}^{(1,8)}$, eq. (37), with $C_{33}^{(S)}$ given by eq. (22) (then a two-time correlation function in the Heisenberg-picture). The

fact that $K_{03300}^{(1,8)}$ and $C_{33}^{(S)}$ are identical means, that a measurement of $K_{03300}^{(1,8)}$ can be used to infer the unperturbed $C_{33}^{(S)}(T_2, T_4)$. This holds, correspondingly, for $K_{00330}^{(1,8)}$ and $K_{00033}^{(1,8)}$. In this sense time-correlations of the past still “coexist”.

As we continue into the cycles $m > 1$, the unique identification of tape state and head history is gradually lost: histories become undecidable. The “meaning” of those measurements thus strongly depends on the step- and cycle number. At the end of cycle $m + p$, to be sure, the original situation is restored. The time-parameters $T_{2\mu}$ labelling those histories are thus defined only modulo pT_8 (if period p exists).

6 Special Machines

6.1 A “Coin-Tossing Machine”

For the machine defined by $\{\alpha_\mu = \pi/2; \mu = 1, 2, 3, 4\}$ all pair correlations and all one-point expectation-values are zero by the end of cycle $m = 1$ (cf. eqs. (37), (36)). The resulting histories all have the same probability and look like those of independent coin tossings at the times $T_{2\mu}$. As for the “Zeno-machine” $\{\alpha_\mu = \pi/M; \mu = 1, 2, \dots, M\}$, a complete measurement of the tape state at the end of cycle $m = 1$ would allow us to reconstruct the history of S . The period is $p = 4$.

6.2 A “Cat Machine”

As a next example let us consider the Turing machine defined by $\{\alpha_1 = \pi/2, \alpha_2 = \alpha_3 = \alpha_4 = 0\}$. The period is $p = 8$, again independent of M : $|\psi^{(m,j)}\rangle = |\psi^{(m+8,j)}\rangle$. At the end of any cycle m all memory cells are strictly correlated (cf. eq. (37)). Furthermore,

$$|\psi^{(1,8)}\rangle = \frac{1}{\sqrt{2}}(|11110\rangle - i|00001\rangle) \quad (42)$$

is found to be a so-called cat-state, for which the decision tree of Fig 4.2 collapses to two histories only, (1111) and (0000), respectively. Moreover $|\psi^{(5,8)}\rangle$ is a different one. As a process the built-up of these cat-states is thus quite simple. While cat states are reduced to product states by the decay (measurement) of any individual subsystem, all the memory pair correlations discussed here remain intact as long as the decaying subsystem is not part of that very pair.

6.3 Large-Scale Predictability

For $m = 100 < p$ and $M + 1 = 10$ we would have roughly $m2^M \approx 5 \cdot 10^4$ transformations in a $2^{M+1} \approx 1000$ -dimensional Hilbert-space; nevertheless, the calculation of these expectation values would scale, at most, linearly with M , independent of m ! This indicates that simulations even of large networks could become feasible based on such rules. Of course, the number of expectation values increases exponentially with the system size $M + 1$.

7 Conclusions

We have discussed the dynamics of a special quantum network, which combines quantum-mechanical and classical features: The quantum-mechanical variables consist of a “Turing head” (pseudospin S) and a “Turing tape” (M memory spins). Classical variables are the phenomenological Hamilton-parameters, which are switched externally to generate discrete unitary transformations. The machine behavior is defined by its initial state and the phase angles α_μ specifying those transformations.

This switching can be visualized as being induced by the Turing head performing pre-determined cycles over $2M$ Turing head positions. Correlations in terms of multi-point expectation values are built up in this process. Time defines the order of non-commuting operations and quantitatively controls transformation parameters.

The structure of these correlations may be attributed to the notorious “holistic nature” of quantum mechanics. Nevertheless, this built-up follows a strict logic: the type of admissible manipulations (rotations) is severely constrained in all but the simplest 2-level-space; this observation certainly applies to our present 2^{M+1} -level-model. Additional constraints are built in by the selection of transformations which are actually implemented. Here they relate to the fact that the multi-levels actually refer to $M + 1$ subsystems. These constraints are reflected by the spatio-temporal pattern of correlations.

There is probably good news and there is bad news as far as the consequences are concerned: The bad news is that the implementation of specific processes is much more constrained in the quantum regime than in the macroscopic world; this makes experimental progress in quantum computation depressively slow.

The good news could be that, eventually, only constrained systems can make up a useful machinery; systems with large, unrestricted state spaces (like a free gas) are “useless”. The constraints are something like fixed axles, wheels, and connecting rods in classical mechanics. Under fairly moderate conditions those correlations and the correlation between correlations should constitute a machine behavior. Rather than *enforcing* some specific behavior defined by abstract algorithms we might be better off trying to exploit the experimental repertoire of *real* quantum networks.

References

1. Ekert, A. and Jozsa, R.: Rev. Mod. Phys. **68** (1996) 1
2. Barenco, A. et. al.: Phys. Rev. A **52** (1995) 3457
3. Cirac, J. I. and Zoller, P.: Phys. Rev. Lett. **74** (1995) 4091
4. Domokos, P., Raimond, J. M., Brune, M. and Haroche, S.: Phys. Rev. A **52** (1995) 3554
5. Gershenfeld, N. A. and Chuang, I. L.: Science **275** (1997) 350
6. Shnirman, A., Schön, G. and Hermon, Z.: Phys. Rev. Lett. **79** (1997) 2371
7. Mahler, G. and Weberruss, V. A.: Quantum Networks: Dynamics of Open Nanostructures, Springer New York (1995); 2nd revised edition (1998)

8. Ferrero, M. and Santos, E.: *Found. Phys.* **27** (1997) 765
9. Schlienz, J. and Mahler, G., *Phys. Rev. A* **52** (1995) 4396
10. Deutsch, D.: *Proc. Roy. Soc. A* **400** (1985) 97
11. Paz, J. P. and Mahler, G.: *Phys. Rev. Lett.* **71** (1993) 3235
12. Obermayer, K., Teich, W. G. and Mahler, G.: *Phys. Rev. B* **37** (1988) 8111
13. Misra, B. and Sudershan, E. C. G.: *J. Math. Phys.* **18** (1977) 756
14. Knight, P.: *Nature* **344** (1990) 493
15. Itano, W. M., Heinzen, D. J., Bollinger, J. J. and Wineland, D. J.: *Phys. Rev. A* **41** (1990) 2295
16. Omnes, R.: *Rev. Mod. Phys.* **64** (1992) 339

Quantum Effects in Algorithms

Richard Jozsa

School of Mathematics and Statistics,
University of Plymouth, Plymouth, Devon PL4 8AA, U.K.
`rjozsa@plymouth.ac.uk`

Abstract. We discuss some seemingly paradoxical yet valid effects of quantum physics in information processing. Firstly, we argue that the act of “doing nothing” on part of an entangled quantum system is a highly non-trivial operation and that it is the essential ingredient underlying the computational speedup in the known quantum algorithms. Secondly, we show that the watched pot effect of quantum measurement theory gives the following novel computational possibility: suppose that we have a quantum computer with an on/off switch, programmed ready to solve a decision problem. Then (in certain circumstances) the mere fact that the computer *would* have given the answer *if* it were run, is enough for us to learn the answer, even though the computer is in fact *not* run.

1 Introduction

Many recent developments in quantum computation are motivated by existing results in theoretical computer science, adapted and rewritten in a quantum context. This includes much of the recent work on quantum error correcting codes (see for example [3,45]) and also the idea of using the Fourier transform to determine periodicity, which underlies many of the known quantum algorithms [1]. There are relatively few results (such as [7]) with no classical analogue, motivated intrinsically from considerations of *physics*. This is a curious situation considering that the entire subject of quantum computation derives from differences between the classical and quantum laws of physics. Apart from the computer science benefits of providing more efficient computation, an important fundamental aspect of the subject is the insight that it might provide for a deeper understanding of the quantum laws and their origins. Computer science and information theory provide an entirely new conceptual framework for considering this question of physics. Thus we will consider the question: what are the essential *physical* effects that give rise to the known computational speedups? And is it possible to use other differences between quantum and classical physics for novel computational possibilities?

2 Quantum Information Processing and Entanglement

It is often said that the power of quantum computation derives from the superposition principle – the ability to do different computations in parallel in

superposition, and combine the results with cleverly arranged interferences. But this explanation is not precise enough because classical waves also exhibit superposition and any effect of superposition can be mimicked by a classical wave system. However there is an essential difference between classical and quantum superposition, which lies in the different way that the two physical theories describe composite systems [2].

Consider n two-level systems. In the classical case we may for example think of each system as comprising the two lowest energy modes of vibration of a string with fixed endpoints together with all superpositions. According to the laws of classical mechanics, the total state space of the composite system is the *Cartesian* product of the n subsystem spaces. Thus no matter how much the strings interact in their physical evolution, the total state is always a product state of the n separate systems. Hence we can say that the information needed to describe the total state grows *linearly* with n (being n times the information needed to describe a single subsystem).

In contrast, according to the laws of quantum mechanics the total space is the *tensor* product of the subsystem spaces and a general state may be written as

$$|\psi_n\rangle = \sum_{i_1, \dots, i_n=0}^1 a_{i_1 \dots i_n} |i_1\rangle \dots |i_n\rangle \quad (1)$$

Thus generally we will have $O(2^n)$ superposition components present and the information needed to describe the total state will grow *exponentially* with n . The novel quantum effect here – the passage from Cartesian to tensor product – is precisely the phenomenon of entanglement i.e. the ability to superpose general product states.

As stated above, quantum entanglement can be readily *mimicked* by classical wave systems: instead of taking n two-level systems, we consider a single classical wave system with 2^n levels, allowing general superpositions of all these levels, and merely interpret these as entangled states via a chosen mathematical isomorphism between $\otimes^n V_2$ and V_{2^n} (where V_k is a k -dimensional vector space). However this mathematical isomorphism is not a valid correspondence for considerations of complexity (i.e. in which we assess the utilisation of physical resources): if the 2^n classical levels are, say, equally spaced energy modes, then to produce a general state in V_{2^n} we will need to invest an amount of energy exponential in n , whereas a general state in $\otimes^n V_2$ will require only a *linear* amount of energy (as at most, each of the n two-level systems will need to be excited). The essential point here is that entanglement allows one to construct exponentially large superpositions with only linear physical resources and this cannot be achieved with classical superposition.

In the sense described above the state $|\psi_n\rangle$ can encode an exponentially large amount of information. This would be of little consequence if we could not process the information in a suitably efficient way. Fortunately the laws of quantum physics allow precisely this possibility, which appears to be at the heart of the computational speedup exhibited by the known quantum algorithms. Suppose we apply a one-qubit gate U to the first qubit of the entangled state

$|\psi_n\rangle$. This would count as just one step of quantum information processing but to compute the result classically (say by matrix multiplication) we would calculate the new amplitudes by

$$\tilde{a}_{j_1 \dots j_n} = \sum_{i=0}^1 U_{j_1 i} a_{i j_2 \dots j_n} \quad (2)$$

where U_{ji} is the unitary matrix for U . Now, this computation involves exponentially many steps: the 2×2 matrix multiplication of U needs to be performed successively 2^{n-1} times for all possible values of the indices $j_2 \dots j_n$. Although the action of U on qubit 1 is a physically simple operation, it is represented mathematically as a tensor product $U \otimes I_2 \otimes \dots \otimes I_2$ (where I_2 is the identity matrix which represents “doing nothing” on qubits 2 to n) and hence mathematically it becomes an exponentially large unitary operation. Thus because of the tensor product rule we can (somewhat enigmatically) state the principle:

(P1): The physical act of doing nothing on part of an entangled composite system is a highly nontrivial operation. It leads to an exponential information processing benefit if used in conjunction with performing an operation on another (small) part of the system.

Indeed it is difficult to process the quantum information by only a “small amount”. Eq. (2) illustrates that any small local operation (addressing a small part of the system) will generally correspond to an exponentially large processing operation from a classical point of view. Intuitively this reflects the denseness of the exponential quantum information stored within the linear resources.

One may object to **(P1)**, claiming that surely the information processing gain arises from the local operation that is actually *performed* (e.g. U above) rather than from the part that is *not* performed (e.g. the $(n-1)$ identity operations above)! To see that this is not the case consider our row of n qubits and suppose now that U operates on the first k qubits (so U is a $2^k \times 2^k$ matrix). Let us compare the number of steps required to perform this transformation in the classical and quantum contexts respectively. It is known that any $d \times d$ unitary matrix may be programmed on a quantum computer in $O(d^2)$ steps [8,13] so the quantum implementation of U will require $O((2^k)^2)$ steps. Classically, direct matrix multiplication for a $d \times d$ matrix requires $O(d^2)$ steps. For U we have $d = 2^k$ and the multiplication must be performed 2^{n-k} times. Thus the classical implementation will require $O((2^k)^2 2^{n-k})$ steps. Hence the ratio of quantum computing effort to classical computing effort is $O(2^k/2^n)$. This ratio decreases if either n is held fixed and k is decreased, or k is held fixed and n is increased. In either case we are increasing the proportion of “doing nothing” and this is giving rise to an increased information processing benefit.

The Fourier transform is a fundamental ingredient [10,17,18] in most of the known quantum algorithms which exhibit a super-classical computational speed-up. This includes the algorithms of Deutsch [10], Simon [11], Shor [12,13] and Grover [15]. Using the mathematical formalism of the fast Fourier transform (FFT) [14], the unitary transformation that is the Fourier transform can be implemented exponentially more efficiently in a quantum context [6] than in

any known classical context. For example, for the group of integers modulo 2^n the classical fast Fourier transform algorithm runs in time $O(n2^n)$ whereas its quantum implementation runs in time $O(n^2)$. An analysis of the implementation of the FFT algorithm in the quantum context, given in detail in [6], shows that the achieved exponential speedup may be entirely attributed to the influence of the principle (P1). This appears to be an essential feature of the speedup exhibited by all known quantum algorithms.

The full (exponentially large) amount of information embodied in the identity of a quantum state $|\psi_n\rangle$ is termed “quantum information”. The formalism of quantum mechanics places an extraordinary limitation on the above entanglement-related benefits of quantum information storage and processing: quantum measurement theory implies severe restrictions on the accessibility of the quantum information in the state. For example, according to Holevo’s theorem [9] we can obtain at most n bits of information about the identity of an unknown state $|\psi_n\rangle$ of n qubits by any physical means whatever. This bound is the same as the information capacity of a classical system with the same number of levels. Thus, curiously, natural physical evolution in quantum physics corresponds to a super-fast processing of (quantum) information at a rate that cannot be matched by any classical means, but then, most of the processed information cannot be read! It is a remarkable fact that these two effects do not annul each other – the small amounts of information that are possible to obtain about the identity of the final processed state do *not* coincide with the particular meagre kinds of information processing that can be achieved by classical computation on the input running for a similar length of time. This disparity directly entails the computational speedup possibilities of quantum computation.

3 Counterfactual Quantum Computation

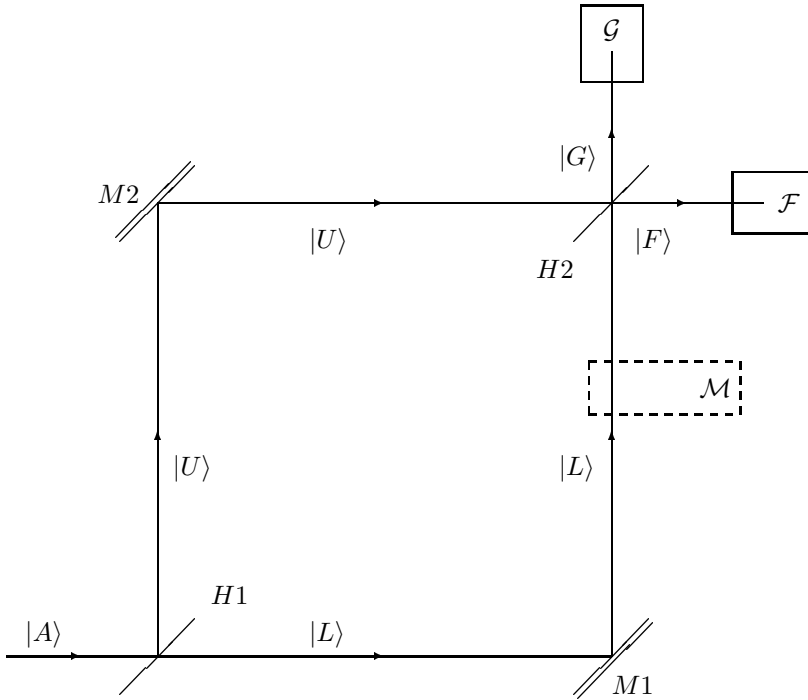
We have argued above that the information processing benefits seen in the known quantum algorithms all rest on some specific features of quantum entanglement. However these features do not exhaust all the ways in which quantum physics differs from classical physics. In an effort to find new quantum algorithms we might ask whether other non-classical features of quantum physics may be exploited for novel computational possibilities (not necessarily just a speedup of computation). Quantum measurement theory (c.f. the inaccessibility of quantum information mentioned above) provides further non-classical aspects of the quantum formalism and these are also related to controversial interpretational issues. We will now describe a novel computational possibility which we call “counterfactual quantum computation”, based on properties of quantum measurement.

A counterfactual effect may be defined as an observable physical effect E whose outcome depends on an event A that might conceivably have happened but in fact did not happen i.e. E is affected by the mere existence of A as a valid possible alternative even though A did not actually occur. Classical physics does not allow physically observable counterfactual effects but quantum physics *does*, at least in the sense described below. Their surprising and somewhat paradoxical

cal occurrence in quantum mechanics has been highlighted in Penrose [20] (see especially §§ 5.2, 5.3, 5.7, 5.8, 5.9, 5.18).

Suppose that we have a quantum computer which has been programmed ready to solve a decision problem. The computer also has an on/off switch, initially set in position off. We will show that in certain circumstances, the mere fact that the computer *would* have given the result of the computation *if* it were run, is sufficient to cause a physically measureable effect from which we can learn the result, even though *the computer is in fact not run!* Our method is based on the so-called Elizur-Vaidman bomb testing problem [21] and the essential idea may be clarified by considering the operation of a simple Mach-Zender interferometer, which we discuss first.

Consider the Mach-Zender interferometer as shown in the following diagram.



Here $H1$ and $H2$ are beam splitters and $M1$ and $M2$ are rigid perfect mirrors. The action of each beamsplitter is taken to be the following (written in terms of the states labelled at $H2$). For horizontal photons

$$|U\rangle \rightarrow \frac{1}{\sqrt{2}}(|F\rangle + |G\rangle) \quad (3)$$

and for vertical photons

$$|L\rangle \rightarrow \frac{1}{\sqrt{2}}(|F\rangle - |G\rangle) \quad (4)$$

A photon enters at $|A\rangle$ and is separated into a superposition $\frac{1}{\sqrt{2}}(|L\rangle + |U\rangle)$ of upper and lower paths. In the absence of the measuring instrument \mathcal{M} the two beams coherently interfere at $H2$ and according to eqs. (3) and (4) the result is $|F\rangle$. Thus the photon is always registered in detector \mathcal{F} and never in detector \mathcal{G} .

Consider now a nondestructive measurement device \mathcal{M} placed in the lower arm, which registers whether or not the photon passed along that arm. The initial state of \mathcal{M} is $|M_0\rangle$ and if a photon is registered it becomes an orthogonal state $|M_1\rangle$. Following the photon we now have

$$|A\rangle \rightarrow \frac{1}{\sqrt{2}}(|U\rangle + |L\rangle) |M_0\rangle \rightarrow \frac{1}{\sqrt{2}}(|U\rangle |M_0\rangle + |L\rangle |M_1\rangle) \quad (5)$$

and the last state may be thought of as the “collapsed” mixture of $|U\rangle |M_0\rangle$ or $|L\rangle |M_1\rangle$, each with probability half. Thus the interference at $H2$ is spoilt and we always have a 50/50 probability of registering the photon in either \mathcal{F} or \mathcal{G} .

Suppose now that the photon is registered in \mathcal{G} and the measurement instrument is seen to be in state $|M_0\rangle$. (This event occurs with probability $\frac{1}{4}$.) Thus the photon has been registered absent in the lower arm and the measurement instrument, having thus apparently done nothing, remains in state $|M_0\rangle$. Yet the photon is seen at \mathcal{G} , which is forbidden in the absence of \mathcal{M} ! Although \mathcal{M} apparently does nothing, it cannot be removed, since then the photon can never register in \mathcal{G} . This is our fundamental counterfactual effect: we can say that the photon can be registered in \mathcal{G} because *if* the photon would have gone along the lower path, it *would* have been detected, even though it did not, in fact, go along the lower arm (since it was not seen by \mathcal{M}).

We can use this effect for computational advantage as follows. Consider an idealised quantum computer which is an isolated physical system containing an on/off switch, a set of program/data registers denoted by the state $|\text{comp}\rangle$ and an output register. The on/off switch is a two-level system with basis states $|\text{on}\rangle$ and $|\text{off}\rangle$ and the output register is a two-level system with basis states $|0\rangle$ and $|1\rangle$. The program/data registers are set up (“programmed”) to solve some given decision problem together with its input (e.g. it might be programmed to test for primality together with a given input integer.) The output register, initially in state $|0\rangle$ will be set by the computation to $|0\rangle$ or $|1\rangle$ according to the answer of the decision problem. The length T of the computation is a known function of the input. The time evolution of the computer for time T is given by

$$\begin{aligned} |\text{on}\rangle |\text{comp}\rangle |0\rangle &\rightarrow |\text{on}\rangle |\text{comp}\rangle |r\rangle \\ |\text{off}\rangle |\text{comp}\rangle |0\rangle &\rightarrow |\text{off}\rangle |\text{comp}\rangle |0\rangle \end{aligned}$$

Here $r = 0$ or 1 is the (initially unknown) result of the computation and the computation will run only if the switch is set to “on”. The result is written into the output register and all program/data registers are returned to their initial state.

Heuristically we will relate this scenario to the interferometer as follows. \mathcal{M} is the quantum computer with $|M_0\rangle$ and $|M_1\rangle$ being the states $|0\rangle$ and $|1\rangle$ of the

output register. The photon is the on/off switch and the two paths are delayed by a time T for the photon to eventually arrive at $H2$. Thus if $r = 0$ the running of the computation makes no distinction between the paths and the photon is always seen in \mathcal{F} . If $r = 1$ the computation (if it ran) would distinguish the two paths and we will see the photon at \mathcal{G} with probability $\frac{1}{2}$. As before, with probability $\frac{1}{4}$ the photon will register at \mathcal{G} (so that we are sure that $r = 1$) and the output register will be seen to be in state $|0\rangle$. Thus the computation has not run, yet we have learnt the result!

More formally in terms of states of the computer, we first set the on/off switch to the superposition:

$$\left(\frac{|\text{off}\rangle + |\text{on}\rangle}{\sqrt{2}} \right) |\text{comp}\rangle |0\rangle \quad (6)$$

and then allow time T (the computation time) to elapse yielding the state

$$\frac{1}{\sqrt{2}} (|\text{off}\rangle |\text{comp}\rangle |0\rangle + |\text{on}\rangle |\text{comp}\rangle |r\rangle) \quad (7)$$

Next rotate the state of the switch by

$$|\text{off}\rangle \rightarrow \frac{1}{\sqrt{2}}(|\text{off}\rangle + |\text{on}\rangle) \quad |\text{on}\rangle \rightarrow \frac{1}{\sqrt{2}}(|\text{off}\rangle - |\text{on}\rangle)$$

This yields the state

$$\begin{aligned} & \frac{1}{\sqrt{2}} \left(\frac{(|\text{off}\rangle + |\text{on}\rangle)}{\sqrt{2}} |\text{comp}\rangle |0\rangle + \frac{(|\text{off}\rangle - |\text{on}\rangle)}{\sqrt{2}} |\text{comp}\rangle |r\rangle \right) \\ &= \frac{1}{\sqrt{2}} \left(|\text{off}\rangle \frac{(|0\rangle + |r\rangle)}{\sqrt{2}} + |\text{on}\rangle \frac{(|0\rangle - |r\rangle)}{\sqrt{2}} \right) |\text{comp}\rangle \end{aligned} \quad (8)$$

Here $r = 0$ or 1 according to the (as yet unknown) result of the computation. Next we measure the switch to see if it is on or off. Note that if $r = 0$ then we never see “on” and if $r = 1$ we see “on” with probability $1/2$. Suppose that we see “on”. Then we know that the result of the computation must certainly be $r = 1$. We then examine the output register which will show $|0\rangle$ with probability $1/2$. If this happens then *the computation has not been run* (because if it had, then the output register must show $|1\rangle$). Overall, if the result is actually $r = 1$ then with probability $1/4$ we learn the correct result (and know it is correct) with no computation having taken place!

Note that if the actual solution of the decision problem is $r = 0$ then we will never ascertain this from the above procedure because if $r = 0$ then the output register will always show 0 and the switch will always be finally seen to be “off”. But this outcome also arises for $r = 1$ with probability $\frac{1}{4}$ and we cannot *a posteriori* distinguish the two possible causes. Correspondingly, if the actual solution is $r = 1$ then with probability $\frac{1}{4}$ will we fail to ascertain this.

The above description of the process represented by eqs. (6) to (8) involves some delicate interpretational issues. For example, a many-worlds adherent might

object that initially the switch was set in an equal superposition of being on and off, so even in the subsequent case of “no computation taking place” the computer actually did run in another “parallel universe” so we cannot claim to get the result for free. One may, to some extent, counter this objection as follows: suppose that when the result is really $r = 1$, the computer is also designed to explode at the end of the computation, if it is run. Then using the above procedure, in *my* world I learn that $r = 1$ and the computer remains *unexploded*, available to do another run. I do not really care if it self-destructs in some “other universe”!

The counterfactual quantum computation procedure above may be considerably improved (using a method inspired by the improvements to the Elizur-Vaidman problem given in [22]) to essentially eliminate the deficiencies noted above. As described below, we will achieve the following:

For any given $\epsilon > 0$

- (i) If the result is $r = 0$, we will learn this with probability 1 but some computation will have taken place.
- (ii) If the result is $r = 1$, we will learn this with probability $1 - \epsilon$ with no computation having taken place.

Thus for the many-worlds adherent, the universe in which the computation takes place can be made to occur with *arbitrarily small* amplitude $O(\sqrt{\epsilon})$ (in the case that $r = 1$), which considerably weakens his/her/its objection. Recall that many basic results in information theory and computer science are formulated in an asymptotic framework which allows an arbitrarily small failure of some desired property. This occurs for example in the distinction between the complexity classes P and BPP [16] (the latter allowing an arbitrarily small probability of a false result) and Shannon’s source coding theorem having not perfect fidelity, but fidelity $1 - \epsilon$ (for any $\epsilon > 0$) for the signals reconstructed from their coded compressed versions. Thus if some undesirable result can be made to occur with arbitrarily small (although non-zero) probability then FAPP it may be ignored. [19]

The improved counterfactual scheme exploits the so-called quantum watched pot effect (or quantum Zeno effect) and it goes as follows. We note first that the state $|\text{comp}\rangle$ will never become entangled with the other registers so we omit it, writing the action of the computer as

$$\begin{aligned} |\text{off}\rangle |0\rangle &\rightarrow |\text{off}\rangle |0\rangle \\ |\text{on}\rangle |0\rangle &\rightarrow |\text{on}\rangle |r\rangle \end{aligned} \tag{9}$$

Choose an angle $\alpha = \frac{\pi}{2N}$ for N sufficiently large (c.f. later). Then perform the following five operations:

- (a) Rotate the switch by angle α .
- (b) Allow the running time T to elapse.
- (c) Read the output register. If it shows 0 then continue. If it shows 1 then discard the state and start again from the beginning.

Remark. (a) and (b) will result in the evolution

$$|\text{off}\rangle|0\rangle \rightarrow (\cos\alpha|\text{off}\rangle + \sin\alpha|\text{on}\rangle)|0\rangle \rightarrow \cos\alpha|\text{off}\rangle|0\rangle + \sin\alpha|\text{on}\rangle|r\rangle \quad (10)$$

If $r = 0$ then the output will *always* show 0 and (c) will result in the state $(\cos\alpha|\text{off}\rangle + \sin\alpha|\text{on}\rangle)|0\rangle$ with probability 1. If $r = 1$ then (c) will result in the collapsed state $|\text{off}\rangle|0\rangle$ obtained with (high) probability $\cos^2\frac{\pi}{2N}$. To complete the procedure we:

- (d) Repeat (a), (b) and (c) a further $N - 1$ times.
- (e) Finally measure the switch to see if it is on or off (assuming that all stages have been kept in (c) and (d)).

We claim that in (e), if the switch is seen to be “on” then r is certainly 0 (and some computation has been done), and if the switch is seen to be “off”, then r is certainly 1 and no computation has taken place. In the latter case the probability of keeping all stages is $(\cos^2\frac{\pi}{2N})^N$ which tends to 1 as $N \rightarrow \infty$. Thus by choosing N to be sufficiently large we can make the probability of success greater than $1 - \epsilon$ for any given ϵ .

To see that our claim is correct, note that if $r = 0$ then the switch is just successively rotated from $|\text{off}\rangle$ to $|\text{on}\rangle$ in N stages and it never entangles with the output register. If $r = 1$ then the state is repeatedly collapsed to $|\text{off}\rangle|0\rangle$ so that no computation takes place in any stage (because if it did, the output register would show the result $r = 1$). Indeed the waiting in (b) acts as a measurement of “on” versus “off” for the switch (if $r = 1$) and in this case, we are just freezing the switch in its $|\text{off}\rangle$ state by frequent repeated measurement. This is the quantum watched pot effect.

Note that according to (i), if $r = 0$ then this result is not learnt “for free”. A natural question is whether or not there is a counterfactual scheme which yields the information of *either* result ($r = 0$ or 1) with no computation having taken place. The procedure described above may be readily modified to provide a scheme with the following properties: with probability $1 - \epsilon$ we learn the result and for either outcome, be it $r = 0$ or $r = 1$, it is obtained for “free” with probability $\frac{1-\epsilon}{2}$. We also learn whether or not the produced result was obtained for “free”. It remains an open question whether or not each of the two results may be obtained for “free” with high probability $1 - \epsilon$, or indeed, whether the sum of these two probabilities can be made to exceed 1.

References

1. Jozsa, R., *Proc. Roy. Soc. Lond. A* **454**, 323–337 (1998)
2. Jozsa, R., “Entanglement and Quantum Computation” in *The Geometric Universe* ed. S. Huggett, L. Mason, K. P. Tod, S. T. Tsou and N. Woodhouse. Oxford University Press. (Also available at <http://xxx.lanl.gov> as quant-ph/9707034.)
3. Steane, A., *Proc. Roy. Soc. Lond. A* **452**, 2551 (1996)
4. Calderbank, A. R. and Shor, P. W., *Phys. Rev. A* **54**, 1098 (1996)
5. Gottesmann, D., *Phys. Rev. A* **54**, 1862 (1996)

6. Ekert, A. and Jozsa, R., "Quantum Algorithms: Entanglement Enhanced Information Processing" appearing in *Phil. Trans. Roy. Soc. Lond.* (1998). (Also available at <http://xxx.lanl.gov> as quant-ph/9803072)
7. Barenco, A., Berthiaume, A., Deutsch, D., Ekert, A., Jozsa, R. and Macchiavello, C., *S.I.A.M. Journal on Computing* **26**, 1541 - 1557 (1997)
8. Deutsch, D., *Proc. Roy. Soc. Lond. A* **400**, 97 (1985)
9. Holevo, A. S., *Probl. Inf. Transm.* **9**, 177 (1973)
10. Deutsch, D. and Jozsa, R., *Proc. Roy. Soc. Lond. A* **439**, 553-558 (1992)
11. Simon, D., *Proc. of 35th Annual Symposium on the Foundations of Computer Science*, (IEEE Computer Society, Los Alamitos), p. 116 (1994) (Extended Abstract). Full version of this paper appears in *S. I. A. M. Journal on Computing* **26**, 1474 (1997).
12. Shor, P., *Proc. of 35th Annual Symposium on the Foundations of Computer Science*, (IEEE Computer Society, Los Alamitos), p. 124 (1994) (Extended Abstract). Full version of this paper appears in *S. I. A. M. Journal on Computing* **26** (Oct 1997) and is also available at LANL quant-ph preprint archive 9508027.
13. Ekert, A. and Jozsa, R., *Rev. Mod. Phys.* **68**, 733 - 753 (1996)
14. Maslen, D. K. and Rockmore, D. N., *Generalised FFT's - a Survey of Some Recent Results*, in *Proc. DIMACS Workshop on Groups and Computation - II* (1995)
15. Grover, L. *Proc. 28th Annual ACM Symposium on the Theory of Computing*, (ACM Press, New York), p. 212-219 (1996)
16. Papadimitriou, C. H., *Computational Complexity* (Addison-Wesley, Reading, MA) (1994)
17. Hoyer, P., "Efficient Quantum Algorithms", preprint available a quant-ph/9702028 (1997)
18. Cleve, R., Ekert, A., Macchiavello, C. and Mosca, M., *Proc. Roy. Soc. Lond. A* **454**, 339-354 (1998)
19. FAPP is an acronym introduced by J. S. Bell meaning "for all practical purposes".
20. Penrose, R.: *Shadows of the Mind*, Oxford University Press (1994)
21. Elitzur, A. C. and Vaidman, L., *Found. of Phys.* **23**, 987-997 (1993)
22. Kwiat, P. G., Weinfurter, H., Herzog, T., Zeilinger, A. and Kasevich, M. A., *Phys. Rev. Lett.* **74**, 4763-4766 (1995)

Automated Design of Quantum Circuits

Colin P. Williams and Alexander G. Gray

Jet Propulsion Laboratory
Mailstop 126-347
4800 Oak Grove Drive
Pasadena, CA 91109-8099

Colin.P.Williams@jpl.nasa.gov, Alexander.G.Gray@jpl.nasa.gov

Abstract. In order to design a quantum circuit that performs a desired quantum computation, it is necessary to find a decomposition of the unitary matrix that represents that computation in terms of a sequence of quantum gate operations. To date, such designs have either been found by hand or by exhaustive enumeration of all possible circuit topologies. In this paper we propose an automated approach to quantum circuit design using search heuristics based on principles abstracted from evolutionary genetics, i.e. using a *genetic programming* algorithm adapted specially for this problem. We demonstrate the method on the task of discovering quantum circuit designs for quantum teleportation. We show that to find a given known circuit design (one which was hand-crafted by a human), the method considers roughly an order of magnitude fewer designs than naive enumeration. In addition, the method finds novel circuit designs superior to those previously known.

1 Introduction: Quantum Circuit Design

1.1 Quantum Computation

Quantum computation is an emerging area of study, which considers the processing of *quantum* information, rather than the familiar classical information. The state of a quantum computer is defined as a superposition of qubits. A computation on such a computer is the unitary evolution of this state, i.e. the action of a unitary matrix operator U upon the state $|\Psi\rangle$. More detailed background on the framework of quantum information processing may be found in [13], [14], and [15].

1.2 Quantum Gates and Circuits

Much recent work has been devoted to the construction of unitary transformations from sequences of more primitive ones. Deutsch ([6]) introduced the notion that such simple unitary operators can be thought of as elementary gates performing logical operations, and more sophisticated operators can be thought of as circuits composed of gates, in analogy to the standard formalism for classical

Boolean electrical circuits. This is sometimes called the network model of computation. Following the classical computation line of analysis, in which certain small sets of gates (as small as one gate) are known to be sufficient to represent all possible circuits, several researchers have proposed such *universal* gate sets (as small as a single parametrized gate family) for quantum circuits ([8], [2]). Besides the identification of such sets, some attempts have been made to characterize the *minimal* number of gates drawn from a given universal set required to implement a given operator U ([7]).

1.3 Circuit Design

Now assume we would like build a circuit to implement a certain computation, represented by U . Most likely our mechanisms for manufacturing quantum computers will begin with allowing us to implement certain very specific primitive quantum operations more effectively than others, for a variety of reasons which will be peculiar to the technology. Given that we have a reasonable set of gates from which to select circuit elements, and perhaps some theoretical ammunition regarding the minimum number we will need, we are still left with the following practical question: What is a specific sequence of those gates that will implement the operation? After we have an efficient and flexible method for answering this question, we will want to answer the following: What is a specific sequence of those gates that will implement the operation using only the minimum number of gates necessary? As the enterprise of building quantum circuits matures, we may eventually wish to find circuits meeting other measures of optimality aside from parsimony. This paper presents a solution to the first (and most important) problem, which also indirectly addresses the issue of parsimony by allowing the size of the circuits considered to vary.

2 Searching the Space of Circuit Designs

2.1 Automated Circuit Design

In this paper we are concerned not with the *theoretical* analysis of minimality of representation, but rather with the practical automated discovery of a correct circuit for a target unitary transformation U . We characterize the problem as a search over the space of possible circuit designs. We focus foremost on demonstrating a search algorithm which finds a correct circuit in less time than it would take to try every possibility. Parsimony of representations will be encouraged through the thoughtful definition of heuristics in the search procedure. It is useful to state here that to avoid exhaustive enumeration, we give up any worst-case guarantee of finding a correct circuit design; so far this is the state of the art in combinatorial optimization ([5]).

2.2 The Search Space

There are two components to a quantum circuit design. One is the topology of the circuit – the gate elements and the connections between them. This is a discrete entity. An important complication enters when we wish to allow topologies to have different sizes, i.e. numbers of gates, which we would prefer to leave unspecified when automating circuit design, leaving the algorithm to find the appropriate size. The second is the assignment of angle values within the gates, if applicable; when our gate selection set includes gates which are actually parametric families of gates, there are continuous parameters to be found.

The paper of DiVincenzo and Smolin ([7]) discussed numerical optimization for the discovery of parameters for two-qubit gates, within a fixed circuit topology, which lead to a desired unitary computation. They used this technique to show that certain gates of interest (the Toffoli gate and arbitrary three-qubit gates) could themselves be represented as circuits of two-qubit gates, by finding the necessary two-qubit gate parameters. In order to find the necessary circuit *topologies*, however, all possible topologies were tried. The focus of that paper was to show the *possibility* of decomposing particular computations into circuits of simpler gates; thus exhaustive enumeration was sufficient as a tool to prove the point. We are interested here in a practical and general method for efficiently finding correct circuit topologies for any given operator, in other words avoiding exhaustive enumeration. We return to the continuous aspect of the search problem later in Section 6.

3 Genetic Programming: A Set of Search Heuristics

3.1 Why Genetic Programming?

Our search problem makes a difficult demand on any search method we might think to employ. First, the search method must be amenable to problems in which it is difficult to characterize the structure of the solution space exactly. To clarify this point, consider that our formulation of the problem leaves the form of the target unitary transformation U completely unspecified; no deep knowledge of U 's substructure, behavior, relationship to the gates used, or nature otherwise can be used to advantage to eliminate invalid possibilities in the search problem. This very general stance is appropriate for quantum circuit design since human techniques and intuitions about quantum circuits have not reached a mature stage yet; once specific classes of quantum circuits can be delineated, it may be fruitful to design search methods which take advantage of their extra constraints. Furthermore, the quantum circuit design problem is one in which it is difficult to evaluate the best next local move to make at any given point in the search; the entire solution must then be evaluated in order to evaluate the effect of a local change in a circuit candidate. Genetic programming is appropriate in this setting since it relies only on evaluations of entire circuits.

Second, it must be capable of considering solution structures of *variable length*. This is crucial if it is to have any hope of finding small designs; it must

be given the latitude to explore solution candidates of different sizes. A particular set of search heuristics, the so-called *genetic programming* method [12], has the distinction of being the *only* search technique having the capability of searching over solutions of varying structure and size. Genetic programming is a type of *genetic algorithm* [9], which in turn is a type of stochastic hill-climbing, or “go with the winners” algorithm ([1]), along with simulated annealing ([10]). Genetic programming is the kind of genetic algorithm which is concerned with non-fixed-length topological structures, rather than the simpler case of fixed-length solutions.

3.2 The Parts of Genetic Programming

Genetic programming is a simple set of search heuristics based loosely on the principles of evolutionary genetics. One of its most distinctive traits is that it is a *population-based* method, or one which maintains multiple solution candidates simultaneously, whose ‘evolution’ paths may interact with each other. In particular, they may trade substructures in an operation called “*crossover*”, in analogy to sexual reproduction. The method is heavily stochastic, sometimes performing random perturbations on solution candidates (“*mutations*”), and greedily selecting the current best solutions to continue pursuing via random sampling weighted by solution quality (“*fitness*”, “*survival of the fittest*”). A typical genetic programming algorithm has this form:

Initialize population with random solutions.

Until the stopping criterion has been reached,

1. Evaluate the quality of each solution in the population.
2. Sample from the population, weighted by solution quality, to form the ‘breeding pool’.
3. For each member of this subset of the population, choose one of the following operations to perform on it:
 - a. Mutation (choose with probability $p(M)$)
 - b. Crossover (choose with probability $p(C)$; requires a partner)

Each iteration of the algorithm is called a “generation”.

Because its directional guidance is based on evaluations of entire solutions, all that is necessary to apply the algorithm to a problem is a well-defined measure of solution quality; it is thus amenable to problems in which it is difficult to evaluate the best local move to make at each partial solution (such as the circuit design problem). The main power of the method, which distinguishes it from simple stochastic local perturbation, is in the crossover operation. If the problem is one in which we expect substructures to contain localized information, i.e. represent meaningful subsolutions (an analogy to subroutines of a program is useful here), then crossover has a hope of successfully transferring a subsolution to a different solution, perhaps increasing its overall quality. In the circuit design problem, it seems reasonable to expect that transferable subcircuits exist. Crossover is also the main mechanism for obtaining topology candidates of different sizes.

4 A Genetic Programming Algorithm for Quantum Circuit Design

For this investigation we designed a genetic programming algorithm tailored specifically for the problem of quantum circuit design.

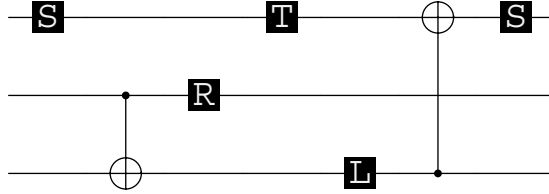


Fig. 1. An example circuit.

4.1 Representation

Circuit Representation. An anonymous quantum circuit is shown in Figure 1 as an example of the representation we use. It is represented as the following nested list data structure, which encodes with each circuit element, its name, parameters if any, and embedding (the wires to which it is connected, followed by the number of wires in the circuit: three in this case):

$$\left[\begin{array}{l} \left\{ \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}, params[], \{1; 3\} \right\}, \\ \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, params[], \{2, 3; 3\} \right\}, \\ \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, params[], \{2; 3\} \right\}, \\ \left\{ \begin{pmatrix} -1 & 0 \\ 0 & -i \end{pmatrix}, params[], \{1; 3\} \right\}, \\ \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, params[], \{3; 3\} \right\}, \\ \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, params[], \{3, 1; 3\} \right\}, \\ \left\{ \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}, params[], \{1; 3\} \right\} \end{array} \right] \quad (1)$$

Gate Selection Set. The algorithm chooses gates from a prespecified selection set. These gates may have unspecified continuous angle parameters associated with them, which must be adjusted by the search algorithm. The gates may also be fixed, or parameterless, gates. In a general setting where little is known about the target transformation, it is sensible to select the gate set such that it forms a universal gate set. It may also be sensible to choose an *overcomplete* set, one which includes a number of gates beyond a computation-universal core subset. This may be useful for obtaining more compact representations, yet may be more costly than having a smaller number of gate types, depending on the technological practicalities of quantum hardware manufacture which hold at the time of the design. An *undercomplete* set may make sense when some known properties of the target computation allow it.

4.2 Evaluation

Solution Quality Measure. To evaluate the quality of a circuit candidate, we compare its matrix form S to the target matrix U using the objective function

$$f(S, U) = \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} |U_{ij} - S_{ij}|, S, U \in U(2^N) \quad (2)$$

This is similar to the objective function used in [7]:

$$f(S, U) = \sum_{i=1}^{2^N} \sum_{j=1}^{2^N} |U_{ij} - S_{ij}|^2, S, U \in U(2^N) \quad (3)$$

We call f the fitness or the *discrepancy*; our goal is to find circuits which minimize the discrepancy between the circuits in our population and the target. When $f = 0$, we have found a circuit which implements U exactly. Otherwise, we have found an approximation to U .

We regard the most sensible evaluation measure as an open question. A paper by Knill [11] considers several measures, many of which are not practically computable, since they take into account all possible states on which the operator may act. One requirement of the measure chosen is that it yields a minimum (maximum) when $S = U$; this property is true of all of Knill's measures. There is a degree of arbitrariness in specifying the proper qualitative behavior of the metric when S differs from U .

While a measure such as f allows the discovery of approximate circuits in a well-defined way, in this paper we focus only upon unitary operations which we can represent *exactly*.

4.3 Selection

Selection is the choosing of a subset from the population to modify in some way. Sampling is weighted by a factor derived from a circuit candidate's discrepancy score, in the way described below, and is performed at the beginning of each generation.

A Ranking-Based Scheme. Rather than translate the discrepancy score of a circuit into its selection probability such that the latter is directly proportional to the score, we instead first order the circuits according to their discrepancies, then determine selection probabilities based directly on the resulting rankings. This procedure has the effect of desensitizing the process with respect to the exact discrepancy distribution, which tends to exhibit extreme ratios between the best candidates and the worst ones; we would like to deemphasize such differences in order to avoid complete domination of the selection process by a few candidates too early in the evolution, which corresponds to entrapment in a local optimum.

Selection Probability Distribution. The circuits are ranked from 1 to N , the number of circuits in the population, 1 denoting the best. Probabilities are defined with which to select members of the population for breeding (i.e. crossover), mutation, and other operations which yield modified solution candidates. We desire a functional form yielding probabilities of selection which decrease as the ranking increases (i.e. gets worse), choosing a quadratic form as a compromise between a form yielding a very weak selection effect (which makes the algorithm closer to a purely random search) such a linear decrease, and a form yielding a very aggressive selection effect (making the algorithm more 'greedy', or susceptible to short-term gains which might cause it to become trapped in a local optimum), such as an exponential decrease.

The probability $P(r)$ of selecting the circuit having ranking r is then $ar^2 + br + c$ for some a , b , and c . To determine some values for these variables we set up some constraints, namely that $P(r)$ is a true probability, i.e. $\sum_{r=1}^N ar^2 + br + c = 1$, that the lowest ranked member is never picked, i.e. $aN^2 + bN + c = 0$, and that the derivative of the probability goes to zero as r goes to N , guaranteeing that the probability function is monotonic decreasing. This set of equations yields values of a , b , and c such that

$$P(r) = \frac{6N}{1 - 3N + 2N^2}r^2 + \frac{6}{N(1 - 3N + 2N^2)}r - \frac{12}{1 - 3N + 2N^2}. \quad (4)$$

To derive the new generation's population from the last generation's members, selection from the described probability distribution is performed N times with replacement; note that the population size stays constant and that on average circuits are multiply represented in the next generation a number of times proportional to their fitness. This process yields the *parents* which are fit enough to draw upon for the various modifications (i.e. search operations) that follow.

To finish the activity of this generation, each parent is replaced by a new circuit resulting from an operation performed on it; the operation to be performed on each circuit is chosen from a discrete probability distribution determined by the user of the algorithm.

4.4 Search Operators

Mutation. Mutation is the random perturbation of a single gate, chosen uniformly at random from the gates within the operand circuit. In the case of fixed gates, i.e. gates without parameters which can vary, the selected gate's embedding is changed by uniformly randomly selecting new connecting lines to replace the old ones.

Substitution. Substitution is similar to mutation, but is the replacement of an existing gate chosen uniformly randomly from the gates within the operand circuit, with another one selected from the gate selection set uniformly randomly. Though replacement can be achieved through an appropriate insertion-deletion pair of operations, described below, its inclusion as a separate operation allows its probability of occurrence to be more explicitly controlled.

Crossover. The circuit resulting from the crossover, or mating, operation is obtained by considering two parent circuits, A and B. A split point is chosen uniformly randomly somewhere along each of the two parent circuits. The circuit resulting from crossover has the first part of the circuit A attached to the second part of the circuit B, or the first part of the circuit B attached to the second part of the circuit A, each with probability 0.5. Note that crossover allows the size of the resulting circuit to change from that of either A or B.

Transposition. Transposition is an operation obtained by generalizing crossover; its result is also defined by considering two parents A and B. A subcircuit is first defined by the selection of beginning and end points in parent A. The beginning point is chosen uniformly randomly along the length of A, and the end point is chosen uniformly randomly from the region between the that point and the end of A. The resulting circuit is found by inserting the subcircuit at a uniformly randomly chosen point along the length of parent circuit B. This also allows the size of the resulting circuit to change from that of either A or B.

Insertion. Insertion is similar to transposition, except that only one parent need be considered; a randomly constructed sequence of gates is inserted at a random point in the parent, resulting in a larger circuit. The beginning and end points of a subcircuit of the parent are chosen as described for the transposition operator, only so that the length of this subcircuit can be used as the length of the random gate sequence to be inserted. This sequence is constructed by choosing uniformly randomly from the gate selection set the described number of gates.

Deletion. Deletion is the inverse of insertion, in that a random subcircuit is chosen from within the parent; this sequence is deleted from the parent, resulting in a smaller circuit.

5 Experimental Results: Quantum Teleportation Circuits

Quantum teleportation has been identified as an important and interesting application of nonlocal effects in quantum mechanics [3]. Brassard has presented a circuit for the 'send' and 'receive' halves of quantum teleportation in [4]. This circuit is compact, requiring only 4 gates in the 'send' subcircuit and 6 in the 'receive' subcircuit. It is shown in Figure 2. The gate definitions can be found in the example circuit shown in [1] and Figure 1.

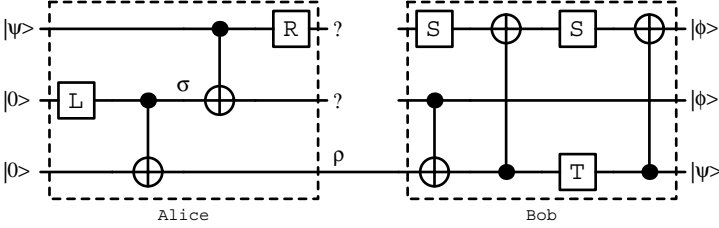


Fig. 2. The quantum teleportation circuit - 'send' and 'receive' parts.

We chose to demonstrate the search algorithm on the computation matrix generated by this circuit, primarily for its general interestingness. Its small size gives the advantage of tractability in the algorithm experimentation phase. Also, because we start with a circuit to obtain the target unitary transform, we know that a compact circuit implementation exists for the problem. We can analyze the computational resources our search method requires to reproduce the hand-designed circuit. As discussed in Section 4.2 using a problem for which an exact circuit representation is known to exist for the gate selection set used avoids the need to consider the appropriateness of the particular fitness measure being used to score inexact circuits.

5.1 The 'Send' Circuit

The algorithm was given the send circuit's computation matrix and a gate selection set consisting of L, R, and XOR. 10 runs were performed, each requiring a different number of generations to find a correct circuit, as follows: 9, 26, 16, 10, 31, 11, 20, 55, 36, 50. 26.4 generations were required on average.

In each case a circuit was found implementing the given computation exactly; although most were different from the original human-designed circuit, all had 4 gates and included at least one each of the L, R, and XOR gates (thus none was necessarily any better than the original circuit). The variance of the number of generations required to find a zero-discrepancy circuit is large, owing to the heavily stochastic nature of the algorithm.

A population size of 100 circuit candidates was used. This is the number of circuit solutions which must be evaluated upon each generation of the algorithm.

Thus, on average, about 2,640 circuits are evaluated for this problem before an answer is found.

By comparison to exhaustive enumeration, the number of possible circuit topologies for this problem, *knowing the number of gates to consider in advance*, can be simply computed as follows: With 3 circuit lines, there are 3 ways to embed the L gate, 3 ways to embed the R gate, and or $\binom{3}{2} = 6$ ways to embed the XOR gate, yielding $3 + 3 + 6 = 12$ different choices for each gate possibility. If we fix the topology size we consider to 4 gates, there are $12^4 = 20,736$ different possible topologies to consider for this problem, using a naive exhaustive approach. Since our search method actually considers circuits of many different sizes, a fair comparison would have to take into account every size class of circuit up to some fairly high number. Our method considered circuits at least as large as 13 gates; note that there are $12^{13} > 10^{14}$ circuits having 13 gates!

We note here that this number does not take into account symmetries and other structure in this search problem, several of which are considered in [7]. Even accounting for these effective reductions of the search space, the computational advantage of a stochastic approach such as the one proposed is still quite significant. Our method may be also be able to take advantage of such information for even greater search efficiency.

Figure 3 shows a typical plot of the average circuit discrepancy over the population at each generation for this problem. The dots on the lower portion of the graph indicate the discrepancy of the best circuit(s) in the population at each generation.

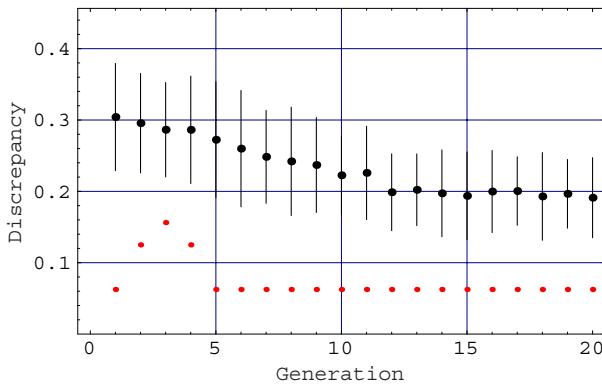


Fig. 3. Typical evolution plot.

5.2 The 'Receive' Circuit

Experiments with the 'receive' part of the circuit demonstrate a further advantage of this approach to automated circuit design beyond achieving a significant

savings in time and computational resources. The flexibility and generality of our approach allows the human user to select a gate set of interest and see whether interesting circuits using those gates are found by the search technique. This type of automated search has the potential to find circuits which are difficult for even resourceful and expert human circuit designers to find. This is true especially when a large number of gates is involved; however this small but practical circuit example illustrates that even modest combinatorial problems are very difficult to find optimal answers for, when unaided by computer methods.

Rather than the original set of gates used in [4] for this circuit, consisting of S, T, and XOR, the genetic programming algorithm was given the gate selection set used above, consisting of L, R, and XOR. One of the resulting exact circuits is shown in Figure 4. Comparing this to the original 'receive' part of the human-designed circuit shown in Figure 2, it is clear that the new circuit is smaller (4 gates versus 6), and that the overall teleportation circuit is more elegant since it requires only 3 types of gates, L, R, and XOR, rather than 5 now that S and T are no longer needed.

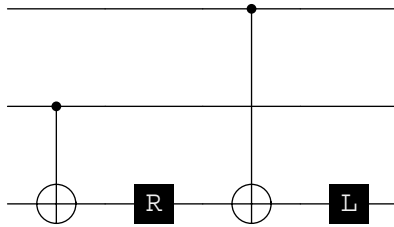


Fig. 4. An efficient circuit found by the search method.

6 Discussion

6.1 Genetic Programming Search as a Tool

At the moment, genetic programming's ability to work with structures of varying sizes makes it the only tool available. Its other primary strength is its effectiveness for opaque problems, where search moves are difficult to evaluate without considering their effect on the entire solution. Rather disappointingly, however, the method's search heuristics are not well-understood formally. For example, issues of convergence, estimated run-time, optimal parameter settings, and behavior dependence on problem context remain empirical issues. Aldous and Vazirani provide one way in which to understand genetic algorithms in general, placing them with simulated annealing in the class of "go with the winners" algorithms ([11]). However, this framework addresses only the 'survival of the fittest' aspect of genetic algorithms, not the effect of the crossover operation, which is one of the hallmarks of genetic algorithms. While much has been written about genetic

algorithms, most analyses have been empirical rather than formal. Genetic *programming*, dealing with variable-length structures, is also surely subsumed by some more general model which can be understood formally – unfortunately this has not yet arrived.

On the positive side, its flexible framework allows the practitioner to plug in his or her own heuristics, encoding any prior knowledge of the problem the user may have (for example, regarding the size of the circuit or the types of gates to use). The specifiable gate selection set allows the specification of only the gates available to the user.

6.2 Extension to Continuous Case

The proposed search method can be extended to allow the inclusion of continuous, or parametrized, gates in the gate selection set, as opposed to the fixed gates used in these experiments. This capability requires necessitates greater computational effort since an optimization must be performed to tune the continuous gate parameters of each circuit candidate such that the discrepancy is minimized given the circuit's discrete topology. However, the ability to incorporate continuous gates holds the promise of more compact circuit solutions, as well as better circuit approximations where necessary. Experiments elucidating this approach, as well as several other potentially powerful extensions, will be described in future reports.

7 Conclusions

In this paper we have formalized the problem of automated quantum circuit design as a search problem. We proceeded to propose a search method tailored for this problem. We then demonstrated its usefulness by showing that it is computationally more efficient than naive enumeration. Finally, we demonstrated that it is capable of discovering useful circuits even when the number of gates considered is small, as exemplified by a novel circuit found by our algorithm for quantum teleportation.

7.1 Acknowledgements

The research described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration and was supported in part by the Center for Integrated Space Microsystems under task number 277-3ROUO-0 and by the JPL Autonomy Program under task number 234-8AX24-0.

References

1. D. Aldous and U. Vazirani, 1997. Go With the Winners Algorithms. UC Berkeley preprint. Available at <http://www.cs.berkeley.edu/~vazirani/>.

2. A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, 1995. Elementary Gates for Quantum Computation, *Phys. Rev. A*, **52**, p. 3457.
3. C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. Wootters, 1993. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, **70**, pp. 1895-1899.
4. G. Brassard, 1996. Teleportation as a quantum computation. In T. Toffoli, M. Bifare, and J. Leao (eds.), *Proceedings of the Fourth Workshop on Physics and Computation* (PhysComp '96), New England Complex Systems Institute, pp. 48-50.
5. T. Cormen, C. Leiserson, R. Rivest, 1993. *Introduction to Algorithms*. Cambridge: The MIT Press.
6. D. Deutsch, 1989. Quantum computational networks. In *Proceedings of the Royal Society of London A*, **425**, p. 73.
7. D. DiVincenzo and J. Smolin, 1994. Results on two-bit gate design for quantum computers. In W. Porod and G. Frazier (eds.), *Proceedings of the Second Workshop on Physics and Computation* (PhysComp '94), IEEE Computer Society Press, pp.14-23.
8. D. DiVincenzo, 1995. Two-bit gates are universal for quantum computation. *Physical Review A*, **51**, pp. 1015-1022.
9. J. H. Holland, 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.
10. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, 1983. Optimization by Simulated Annealing. *Science*, **220**, pp. 671-680.
11. E. Knill, 1995. Approximation by quantum circuits. LANL Report LAUR-95-2225.
12. J. R. Koza, 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge: MIT Press.
13. S. Lloyd, 1993. A potentially realizable quantum computer. *Science*, **261**, pp. 1569-1571.
14. A. Steane, 1997. Quantum computing. Review for *Reports on Progress in Physics*. Available at <http://xxx.lanl.gov/archive/quant-ph/9708022>.
15. C. P. Williams and S.H. Clearwater, 1998. *Explorations in Quantum Computing* (Book plus CD-ROM). Santa Clara: TELOS/Springer-Verlag.

Quantum Search on Structured Problems

Lov K. Grover

3C-404A Bell Labs, 600 Mountain Avenue, Murray Hill NJ 07974
email: lkgrover@bell-labs.com

Abstract. This paper shows how a basic property of unitary transformations can be used for meaningful computations. This approach immediately leads to search-type applications, where it improves the number of steps by a square-root - a simple minded search that takes N steps, can be improved to approximately \sqrt{N} steps. The quantum search algorithm is one of several immediate consequences of this framework. Several novel search-related applications are presented.

1. Introduction

Several interesting problems in computer science can be looked upon as search problems. There are two categories of such problems. First, where the search depends on data in memory - this is the database search kind of problems. Alternatively, the search could be based on a function known in advance - many NP-complete problems and cryptography problems can be expressed in this form. For example the SAT problem of NP-completeness asks whether there exists a combination of binary variables that satisfies a specified set of Boolean equations - this can be looked upon as a search of the state space of the binary variables. In cryptography, the well-known 56-bit DES code (Data Encryption Standard) can be cracked by an exhaustive search of 2^{56} items [BBHT][Phone].

It aroused considerable interest when it was shown that it was possible to improve upon the obvious classical bound for exhaustive search by resorting to quantum mechanics [Search][BBHT] - the intuitive reason for this improvement was that quantum mechanical systems can be in multiple states and simultaneously explore different regions of configuration space. This improved the number of steps by a square-root, i.e. a simple minded search that takes N steps, could be improved to approximately \sqrt{N} steps. The quantum search algorithm was derived using the Walsh Hadamard (W-H) transform and it appeared to be a consequence of the special properties of this transform. Subsequently [Gensrch] showed that similar results are obtained by substituting *any* unitary transformation in place of the W-H transform. This means that a variety of unitary transformations could be used in place of the W-H transform and this leads to algorithms for several different problems. This paper

describes the approach of [Gensrch] and shows how it can be extended to solve various structured problems.

2. Framework

A function $f(x)$, $x=0,1,...(N-1)$, is given which is known to be zero for all x except the single point ($x=t$), the goal is to find t (t for target). The obvious classical technique of searching by looking at the N values of x , one by one, would clearly take $O(N)$ steps.

Assume that we have at our disposal a unitary transformation U that acts on a system with N basis states. First map each value of x to a basis state and start with the system in the basis state s (s for start). If we apply U to s , the amplitude of reaching t is U_{ts} , and if we were to make a measurement that projects the system into a unique basis state, the probability of getting the right basis state would be

$|U_{ts}|^2$. It would, therefore, take $\Omega\left(\frac{1}{|U_{ts}|^2}\right)$ repetitions of this experiment before a

single success. This section shows how it is possible to reach state t in only $O\left(\frac{1}{|U_{ts}|}\right)$ steps. This leads to a sizable improvement in the number of steps if $|U_{ts}| \ll 1$.

Denote the unitary operation that inverts the amplitude in a single state x by I_x . In matrix notation this is the diagonal matrix with all diagonal terms equal to 1, except the xx term which equals -1 .

v_x denotes the column vector which has all terms zero, except for the x^{th} term which is unity.

Consider the following unitary operator: $Q = -I_s U^{-1} I_t U$, since U is unitary U^{-1} is equal to the *adjoint*, i.e. the complex conjugate of the transpose of U . We first show that Q preserves the two dimensional vector space spanned by the two vectors: v_s and $(U^{-1} v_t)$ (note that in the situation of interest, when U_{ts} is small, these two vectors are almost orthogonal).

First consider $Q v_s$. By the definition of Q , this is: $-I_s U^{-1} I_t U v_s$. Note that $v_x v_x^T$ is an $N \times N$ square matrix all of whose terms are zero, except the xx term which is 1. Therefore $I_t = I - 2v_t v_t^T$ & $I_s = I - 2v_s v_s^T$, it follows:

$$(1) \quad \begin{aligned} Q v_s &= -(I - 2v_s v_s^T) U^{-1} (I - 2v_t v_t^T) U v_s \\ &= -(I - 2v_s v_s^T) U^{-1} U v_s + 2(I - 2v_s v_s^T) U^{-1} (v_t v_t^T) U v_s \end{aligned}$$

Using the facts: $U^{-1}U$ and $v_s^T v_s \equiv 1$, it follows that:

$$(2) \quad Qv_s = v_s + 2(I - 2v_s v_s^T)U^{-1}(v_t v_t^T)Uv_s.$$

Simplify the second term of (2) by the following identities: $v_t^T Uv_s \equiv U_{ts}$ and since U is unitary, its inverse is equal to its *adjoint* (the complex conjugate of the transpose) $v_s^T U^{-1}v_t \equiv U_{ts}^*$.

$$(3) \quad Qv_s = v_s \left(1 - 4|U_{ts}|^2\right) + 2U_{ts}(U^{-1}v_t)$$

Next consider the action of the operator Q on the vector $U^{-1}v_t$. Using the definition of Q (i.e. $Q \equiv -I_s U^{-1} I_t U$) and carrying out the algebra as in the computation of Qv_s above, this yields:

$$(4) \quad Q(U^{-1}v_t) \equiv -I_s U^{-1} I_t U(U^{-1}v_t) = -I_s U^{-1} I_t v_t = I_s U^{-1} v_t.$$

Writing I_s as $I - 2v_s v_s^T$ and as in (3), $v_s^T U^{-1}v_t \equiv U_{ts}^*$:

$$(5) \quad Q(U^{-1}v_t) = (U^{-1}v_t) - 2v_s v_s^T (U^{-1}v_t) = (U^{-1}v_t) - 2U_{ts}^* v_s.$$

It follows that Q transforms any superposition of v_s and $(U^{-1}v_t)$ into another superposition of the two vectors, thus preserving the two dimensional vector space spanned by v_s and $(U^{-1}v_t)$. (3) & (5) may be written as:

$$(6) \quad Q \begin{bmatrix} v_s \\ U^{-1}v_t \end{bmatrix} = \begin{bmatrix} (1 - 4|U_{ts}|^2) & 2U_{ts} \\ -2U_{ts}^* & 1 \end{bmatrix} \begin{bmatrix} v_s \\ U^{-1}v_t \end{bmatrix}$$

It follows as in [BBHT], that if we start with v_s , then after η repetitions of Q we get the superposition $a_s v_s + a_t (U^{-1}v_t)$ where $a_s \equiv \cos(2\eta|U_{ts}|)$ & $a_t \equiv \sin(2\eta|U_{ts}|)$ if $|U_{ts}| \ll 1$. If $\eta = \frac{\pi}{4|U_{ts}|}$, then we get the superposition $U^{-1}v_t$; from this with a single

application of U we can get v_t . Therefore in $O\left(\frac{1}{|U_{ts}|}\right)$ steps, we can start with the s -state and reach the target state t with certainty.

3. Quantum Operations

The interesting feature of the analysis of section 2 is that U can be *any* unitary transformation, whatsoever. Clearly, it can be used to design algorithms where U is a transformation on the qubits in a quantum computer - the object of this paper is to present some such applications. Quantum mechanical operations that can be carried out in a controlled way are unitary operations that act on a small number of qubits in each step. It is possible to design a variety of quantum mechanical algorithms using

just a few elementary quantum mechanical operations. Two of the elementary unitary operations needed are: the W-H transformation operation and the selective inversion of the amplitudes of certain states.

2^n

A basic single bit operation in quantum computing is the operation M - this is represented by the following matrix: $M \equiv \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, i.e. a bit in the state 0 is

transformed into a superposition: $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$. Similarly a bit in state 1 is transformed

into $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$. In a system in which the states are described by n bits (it has

$N = 2^n$ possible states) we can perform the operation M on each bit independently in sequence thus changing the state of the system. The state transition matrix representing this operation will be of dimension $2^n \times 2^n$. Consider a case when the starting state is one of the 2^n basis states, i.e. a state described by an n -bit binary string with some 0s and some 1s. The result of performing the operation M on each bit will be a superposition of states described by all possible n -bit binary strings with the

amplitude of each state being $\pm 2^{-\frac{n}{2}}$. To deduce the sign, observe that from the definition of the matrix M , i.e. $M \equiv \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, that the phase of the resulting

configuration is changed only when a bit that was previously a 1 remains a 1 after the transformation is performed. Hence if \bar{x} be the n -bit binary string describing the starting state and \bar{y} the n -bit binary string describing the resulting string, the sign of the amplitude of \bar{y} is determined by the parity of the bitwise dot product of \bar{x} and \bar{y} , i.e. $(-1)^{\bar{x} \cdot \bar{y}}$. This transformation is the W-H transformation [DJ]. This operation

(or a closely related operation called the Fourier Transformation [Factor]) is one of the things that makes quantum mechanical algorithms more powerful than classical algorithms and forms the basis for most significant quantum mechanical algorithms.

The other transformation we will need is the selective inversion of the phase of the amplitude in certain states. Unlike the W-H transformation and other state transition matrices, the probability in each state stays the same since the square of the absolute value of the amplitude in each state stays the same. The following is a realization based on [BBHT]. Assume that there is a binary function $f(x)$ that is either 0 or 1. Given a superposition over states x , it is possible to design a quantum circuit that will selectively invert the amplitudes in all states where $f(x)=1$. This is achieved by appending an ancilla bit, b and considering the quantum circuit that transforms a state $|x, b\rangle$ into $|x, f(x)XOR b\rangle$ (such a circuit exists since, as proved in [Revers], it is possible to design a quantum mechanical circuit to evaluate any function $f(x)$ that can be evaluated classically). If the bit b is initially placed in a superposition

$\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, this circuit will invert the amplitudes precisely in the states for which $f(x)=1$, while leaving amplitudes in other states unchanged.

4. Summary of Applications

As mentioned in section 1, the search problem is the following: a function $f(x)$, $x=0,1,\dots,(N-1)$, is given which is known to be non-zero at certain values of x ; the task is to find one such value. No structure is known for $f(x)$ except for what is explicitly mentioned in the specific problems (4.1)...(4.7). N is assumed to be a power of 2, say $N=2^n$. There is a one-to-one correspondence between the N values of x and the respective states of an n -bit register. States corresponding to values of x for which $f(x)$ is non-zero, are referred to as t -states.

- (4.1) Assume that $f(x)=0$ everywhere except for a single value of x . This is the standard problem of exhaustive search.
- (4.2) As in (4.1), there is a single point, t , where $f(x)$ is non-zero. Some information about t is available in the following form - another n bit word, r , is given which is known to differ from t in at most k out of the n bits.
- (4.3) There are multiple points (t -states) at which $f(x)$ is non-zero, it is required to find any one of these. Some structure to the problem is specified in the following form. We are given a certain unitary transformation U & multiple s -states so that U_{ts} for any s & any t are the same. The setting of this subsection is abstract, (4.4), (4.5) & (4.7) apply the framework of this subsection to concrete problems.
- (4.4) As in (4.3), there are multiple points at which $f(x)$ is non-zero, it is required to find any one of these. However, unlike (4.3), no further structure to the problem is given.
- (4.5) There is a single point t where $f(x)$ is non-zero. Some information about t is available in the form of l n -bit strings, each of which differ from t in exactly k out of n bits.
- (4.6) $f(x)=0$ everywhere except at the unique point $x=t$, it is required to find t . Also, as in (4.3), we are given a unitary transformation U and multiple s -states. However, unlike (4.3), U_{ts} for various s -states & various t -states are not all identical.

The analysis of section 2 extends the power of quantum search so that it can be used with an arbitrary unitary transform U , but only with a single s and single t -state. (4.3) extends it to multiple states, but in a restricted way. This derivation extends to multiple s -states. It is still not known how to handle multiple t -states that are not exactly symmetric.

(4.7) This problem illustrates how the abstract techniques discussed earlier can be applied to solve an actual problem. This problem was first discussed by Eddie Farhi & Sam Gutmann [Structure].

Two functions $f(x, y)$ & $g(x)$ are defined on the domain $x = 0, 1, \dots, (N-1), y = 0, 1, \dots, (N-1)$, . $f(x, y)$ is zero everywhere except at the unique point (t_1, t_2) , $g(x)$ is non-zero at M values of x including $x = t_1$ (here $M \ll N$). The problem is to find t_1 & t_2 . Classically this problem would take $\Omega(NM)$ steps. The algorithm of (4.1), without using the function $g(x)$, would take $O(N)$ steps. The following analysis makes use of the general technique of (4.3) to develop an $O(\sqrt{NM})$ step algorithm. Several variants of this problem are also considered.

4.0 The Approach

The following general approach is made use of in each of the next 7 sub-sections – (4.1)...(4.7).

There are $N = 2^n$ states, represented by n qubits, the task is to get the system into some target state(s) t at which $f(x)$ is non-zero. A unitary transform U and the initial state s are selected and U_{ts} is calculated. It then follows by section 2 that by $\frac{\pi}{4|U_{ts}|}$ repetitions of the operation sequence $-I_s U^{-1} I_t U$, followed by a single application of U , the initial state s is transformed into the final state t .

4.1 Exhaustive Search

Assume that $f(x) = 0$ everywhere except at a single point t . The object is to find t .

As mentioned in the first paragraph of the introduction, there are several important problems in computer science for which there no solution is known, except exhaustive search.

Solution For the W-H transform, described in section 3, U_{ts} between any pair of states s & t is $\pm \frac{1}{\sqrt{N}}$. Therefore we can start with any state s and the procedure of

(4.0) gives us an algorithm requiring $O\left(\frac{1}{|U_{ts}|}\right)$ steps, i.e. $O(\sqrt{N})$ steps.

In case s be chosen to be the $\bar{0}$ state, then the operation sequence $Q = -I_{\bar{0}} W I_t W$ leads to the standard quantum search algorithm based on the *inversion about average* interpretation [Gensrch] (note that $W^{-1} = W$). To see this write $I_{\bar{0}}$ as $I - 2v_{\bar{0}}v_{\bar{0}}^T$.

Therefore for any vector \bar{x} : $-WI_0W\bar{x} = -W\left(I - 2v_0v_0^T\right)W\bar{x} = -\bar{x} + 2Wv_0v_0^TW\bar{x} : U_{ts}$.

It is easily seen that $Wv_0v_0^TW\bar{x}$ is another vector each of whose components is the

same and equal to A where $A \equiv \frac{1}{N} \sum_{i=0}^{N-1} x_i$ (the average value of all components).

Therefore the i^{th} component of $-WI_0W\bar{x}$ is simply: $(x_i + 2A)$. This may be written as $A + (A - x_i)$, i.e. each component is as much above (below) the average as it was initially below (above) the average, which is precisely the *inversion about average*.

In case s be chosen to be a state different from $\bar{0}$, the dynamics is still very similar to the standard quantum search algorithm; however, the *inversion about average* interpretation no longer applies.

4.2 Search when an Item *near* the Desired State Is Known:

This problem is similar to (4.1), i.e. $f(x)=0$ except at the single point t . The difference from (4.1) is that some information about the solution, t , is available in the following form: another n bit word, r , is specified - t is known to differ from r in at most k of the n bits.

Such a problem would occur in any situation when we had some prior information about the solution, this information could come either from prior knowledge or from a noisy data-transmission.

Solution: The effect of the constraint is to reduce the size of the solution space. One way of making use of this constraint, would be to map this to another problem and then exhaustively search the reduced space using (4.1). However, such a mapping would involve additional overhead. This section presents a different approach which carries over to more complicated situations as in (4.5).

Instead of choosing U as the W-H transform, as in (4.1), in this section U is tailored to the problem under consideration. The starting state s is chosen to be the specified word r . The operation U consists of the following unitary transformation

$$\begin{bmatrix} \sqrt{1-\frac{k}{n}} & \sqrt{\frac{k}{n}} \\ \sqrt{\frac{k}{n}} & \sqrt{1-\frac{k}{n}} \end{bmatrix}$$
, applied to each of the n qubits. Calculating U_{ts} , it follows that

$|U_{ts}| = \left(1 - \frac{k}{n}\right)^{\frac{n-k}{2}} \left(\frac{k}{n}\right)^{\frac{k}{2}}$ and $|U_{ts}| = \frac{n}{2} \log \frac{n-k}{n} - \frac{k}{n} \log \frac{n-k}{k}$. The technique

described in (4.0) can now be used - as in (4.1), this consists of repeating the sequence

of operations , $-I_s U I_t U$, $O\left(\frac{1}{|U_{ts}|}\right)$ times, followed by a single application of the operation U (note that, as in (4.1), $U^{-1} = U$).

The size of the space being searched in this problem is approximately $\binom{n}{k}$ which is equal to $\frac{n!}{n-k!k!}$. Using Stirling's approximation: $\log n! \approx n \log n - n$, it follows that $\log \binom{n}{k} \approx n \log \frac{n}{n-k} - k \log \frac{k}{n-k}$, comparing this to the number of steps required by the algorithm, we find that the number of steps in this algorithm, as in (4.1), varies as the square-root of the size of the solution space being searched.

4.3 Multiple s & t States with the Same U_{ts} :

$f(x)$ is non-zero at β values of x , i.e. there are β t -states. Some structure of the problem is specified in the following form. Assume that we have at our disposal a unitary transform U and α s -states such that U_{ts} between *any* t -state and *any* s -state is the same. The object is to find one of the t -states. This is accomplished by transforming the system into a superposition so that there is an equal amplitude in each of the t -states and zero amplitude elsewhere. After this, a measurement is made that projects the system into one of its basis states, this gives a t -state.

The problem considered in this subsection is abstract in the sense U is an arbitrary unitary transformation. (4.4), (4.5) and (4.7) apply this to concrete problems.

Solution: The approach is similar to the exhaustive search problem of (4.1). However, the analysis of section 2 has to be redone with the following three changes:

- (a) The starting state instead of being v_s , is the superposition $\frac{1}{\sqrt{\alpha}} \sum_{a=0}^{\alpha-1} v_{s_a}$ - the

amplitude in all s_a states is equal to $\frac{1}{\sqrt{\alpha}}$, and zero everywhere else. Assuming α to

be a power of 2 ($\alpha \equiv 2^a$), such a superposition can be easily created by the following procedure. Start with an a bit system with all bits in the 0 state. Do a W-H transform on the a bit system and then carry out a mapping from the 2^a states to the s -states.

(b) The operations I_s & I_t invert the amplitudes in *all* s -states & *all* t -states, respectively.

(c) It can then be shown by an analysis similar to section 2, that after

$O\left(\frac{1}{\sqrt{\alpha\beta}|U_{ts}|}\right)$ operations of $-I_s U^{-1} I_t U$ followed by a single application of U , the

system reaches a superposition so that the amplitude is equal in all t -states and is zero everywhere else. Note that the number of operations is smaller by a factor as compared to the situation when there were single s -states & single t -states (as in (4.1)).

4.4 Problem

$f(x)$ is non-zero at β values of x , equivalently there are β t -states - the task is to find one of these.

This is the problem of exhaustive search when there are multiple (β) solutions. A classical search would take an average of $O\left(\frac{N}{\beta}\right)$ steps to find a solution.

This section presents an $O\left(\sqrt{\frac{N}{\beta}}\right)$ step quantum mechanical algorithm.

Solution By the definition of the W-H transform in section 3, $W_t\bar{0}$ for any t is the same. Therefore if we choose s as the $\bar{0}$ state, then it follows by (4.3) that after

$O\left(\sqrt{\frac{N}{\beta}}\right)$ repetitions of $-I_s W I_t W$ followed by a single application of W , the system reaches a superposition such that the amplitude is equal in all the t states and zero everywhere else. Note the following three points regarding this scheme:

- As in (4.3), the operation inverts the phase for all β t -states.
- The above implementation requires β to be known in advance.
- The search time is $\sqrt{\beta}$ faster than the exhaustive search algorithm of (4.1).
- It is necessary to choose s as the $\bar{0}$ state, this is different from (4.1) where s could be arbitrary.

4.5 Problem

$f(x)=0$ except at the single point t . Some information about t is available in the form of α n -bit strings, each of which differs from t in *exactly* k bits.

This is in some sense the dual of (4.4). In that case there were multiple t -states but a single s -state, while in this problem there are multiple s -states and a single t -state. This kind of problem could occur in extracting a signal out of multiple noisy transmissions.

Solution Let the α specified states be the s -states. Initialize the system to a superposition of these states by the process described in (4.3)(a). After this, apply the

unitary transform U which applies the following unitary operation

$$\begin{bmatrix} \sqrt{1-\frac{k}{n}} & \sqrt{\frac{k}{n}} \\ \sqrt{\frac{k}{n}} & \sqrt{1-\frac{k}{n}} \end{bmatrix}$$

to each qubit. As in (4.2), $|U_{ts}| = \left(1 - \frac{k}{n}\right)^{\frac{n-k}{2}} \left(\frac{k}{n}\right)^{\frac{k}{2}}$ and

$|U_{ts}| = \frac{n}{2} \log \frac{n-k}{n} - \frac{k}{n} \log \frac{n-k}{k}$ for all s -states. Also, since each of the s -states differ from t in exactly the same number of bits implies that U_{ts} has the same sign for all s -states. The framework of (4.3) can now be used - this yields an algorithm that is $\sqrt{\alpha}$ times faster than that of (4.2).

4.6 Multiple s -States & a Single t -State Such that U_{ts} between Various s -States & t Is Not Identical

$f(x) = 0$ everywhere, except at the single point t . As in (4.3), some structure to the problem is specified via U . Assume that we have at our disposal a unitary transform U and various s -states are specified. However, unlike (4.3), U_{ts} between various s and various t -states are *not* exactly equal. This case is qualitatively different from all of those considered so far because the analysis of section 2 or the modified analysis of (4.3), does not directly apply.

This extends the noisy data-transmission problem of (4.5), to the case where there are multiple states specified that are close to the solution state but differ from it in varying number of bits ((4.5) required each of the given states to differ from the solution in *exactly* the same number of bits).

Solution Assume the number of s -states to be α , further assume that α is a power of 2, i.e. $\alpha \equiv 2^a$. Consider V to be a unitary matrix which is a product of 3 simpler unitary matrices, i.e. $V \equiv V_1 V_2 U$ and $V^{-1} \equiv U^{-1} V_2^{-1} V_1^{-1}$.

V_1 is a W-H transformation on a bits and α states where $\alpha \equiv 2^a$,

V_2 maps the $\alpha \equiv 2^a$ states generated by V_1 onto the respective s -states in N -dimensional state space,

U is the available unitary transform on the N states.

Let the initial state be the $\bar{0}$ state. As a result of $V_1 V_2$, the amplitude in each of the s_a states: $s_0, s_1 \dots s_{\alpha-1}$, becomes $\frac{1}{\sqrt{\alpha}}$; after U , the amplitude in the t -state is

$\frac{1}{\sqrt{\alpha}} \sum_{a=0}^{\alpha-1} U_{ts_a}$. By (4.0), it follows that after $\left\lceil \frac{\sqrt{\alpha}}{\sum_{a=0}^{\alpha-1} U_{ts_a}} \right\rceil$ repetitions of $Q = -I_0 V^{-1} I_t V$

followed by a single application of V , the amplitude in the target state becomes $O(1)$.

Many of the problems in the previous subsections can be seen to be particular cases of this. For example in case U_{ts_a} are all equal, say to u , then the number of iterations

becomes: $\frac{1}{|u|\sqrt{\alpha}}$. The algorithm and bound of (4.5) immediately follow from this.

4.7 Two Dimensional Search

Two functions $f(x, y)$ & $g(x)$ are defined on the domain $x = 0, 1, \dots, (N-1)$, $y = 0, 1, \dots, (N-1)$. $f(x, y)$ is zero everywhere except at the unique point (t_1, t_2) , $g(x)$ is non-zero at M values of x including $x = t_1$ (here $M \ll N$). The problem is to find t_1 & t_2 .

Classically this problem would take $\Omega(NM)$ steps. The algorithm of (4.1), without using the function $g(x)$ would take $O(N)$ steps. This section presents an $O(\sqrt{NM})$ step algorithm. For a different analysis, along with a proof that the algorithm of this section is within a constant factor of the fastest possible algorithm, see [Structure].

Several variations of this problem are also briefly considered, these demonstrate how the techniques discussed in this paper can be applied to real problems.

Solution First consider the function $g(x)$, $x = 0, 1, \dots, (N-1)$. By executing the algorithm of (4.4), with the t -states as the non-zero values of $g(x)$, it is possible for the system to reach a superposition such that at each point at which $g(x)$ is non-zero, the amplitude is $\frac{1}{\sqrt{M}}$ and the amplitude is zero everywhere else. This is

accomplished by a sequence of $O\left(\sqrt{\frac{N}{M}}\right)$ unitary transformations (4.4) - denote this

composite unitary operation by U_1 . Next keep the value of x the same and carry out the algorithm of (4.1) on the N values of y with the t -state corresponding to the non-zero value of the function $f(x, y)$. This consists of a sequence of $O(\sqrt{N})$ elementary unitary transformations, denote this composite unitary operation by U_2 . It follows from (4.1), that as a result of this operation sequence, in case the particle is at $x = t_1$, it is also at $y = t_2$. The amplitude of the system being in the desired state is therefore

$\frac{1}{\sqrt{M}}$. By means of the unitary transformation $U = U_1 U_2$, the system starting from a certain initial state reaches the desired state with an amplitude of $\frac{1}{\sqrt{M}}$.

Since U is a sequence of elementary unitary operations, it follows that U^{-1} is a sequence of the adjoints of the same operations in the opposite order and can hence be synthesized. By applying the procedure of section 2, it follows that by repeating the sequence of operations $Q \equiv -I_s U^{-1} I_t U$, $O(\sqrt{M})$ times, the system reaches the target state with certainty.

The total number of steps is given by the number of repetitions of Q (i.e. $O(\sqrt{M})$) times the number of steps required for each repetition, (i.e. $O\left(\sqrt{\frac{N}{M}}\right) + O(\sqrt{N})$) which gives $O(\sqrt{M}) \times \left(O\left(\sqrt{\frac{N}{M}}\right) + O(\sqrt{N})\right) = O(\sqrt{NM})$.

The above analysis easily extends to more general cases. For example consider the case, where the search-space is rectangular instead of square, i.e. the number of possible values for x is N_1 and for y is N_2 . The number of steps, instead of being $O(\sqrt{NM})$, now becomes $O(\sqrt{N_1}) + O(\sqrt{N_2 M})$. Alternatively consider the case where there is an η dimensional space with the same number of points (N) in each dimension. Instead of just $f(x, y)$ & $g(x)$, there are now η functions: $f(x_1, x_2, \dots, x_\eta)$, $g_1(x_1, x_2, \dots, x_{\eta-1})$, $g_2(x_1, x_2, \dots, x_{\eta-2})$, ..., $g_{\eta-1}(x_1)$ with analogous definitions. It follows by a similar approach that the number of steps is now $O(\sqrt{NM_1 M_2 \dots M_{\eta-1}})$.

Another variation is when the function $f(x, y)$ is non-zero at multiple points, say β points, and it is required to find one of these. In case each of the β points has a different value of x , the framework of (4.3) applies. Considering the unitary transformation U as $U_1 U_2$, leads to an algorithm that requires $O\left(\sqrt{\frac{NM}{\beta}}\right)$ steps.

It is not clear how to derive an algorithm when some of the β points have the same value of x .

5. Conclusion & Further Work

[Search] demonstrated how to make use of quantum mechanical properties to develop exhaustive search kinds of algorithms, i.e. algorithms for problems that

lacked any structure. [Search] used subtle properties of a particular quantum operation called the W-H transform. Subsequently [Gensrch] extended this so that other quantum operations could be used instead of the W-H transform.

Most interesting problems in computer science are concerned with the structure of problems and how to develop algorithms to take advantage of this structure. [Hogg] has previously suggested heuristic quantum mechanical algorithms for structured problems. This paper has given several examples of structured problems and how search-type algorithms can be extended to solve these ((4.2) through (4.7)) - quantitative closed form bounds were derived for the running time of these algorithms. The extensions have shown how to deal with the situation where there are multiple s -states (initial states) and a single t -state (target state) ((4.5) and (4.6)). Also, it is possible to deal with multiple t -states provided they are exactly symmetric ((4.3) and (4.4)). The next step would be to obtain a general algorithm with multiple t -states.

6. References

- [BBHT] M. Boyer, G. Brassard, P. Hoyer & A. Tapp, *Tight bounds on quantum searching*, Proc., PhysComp 1996 (lanl e-print quant-ph/9605034).
- [BV] E. Bernstein and U. Vazirani, *Quantum complexity theory*, Proc. 25th ACM Symposium on Theory of Computing, 1993; for full paper see - SIAM Journal on Computing, vol. 26, no. 5, Oct 1997.
- [BBBV] C. H. Bennett, E. Bernstein, G. Brassard & U. Vazirani, *Strengths and weaknesses of quantum computing*, SIAM Journal on Computing, pp. 1510-1524, vol. 26, no. 5, October 1997.
- [DJ] D. Deutsch & R. Josza, *Rapid solution of problems by quantum computation*, Proc. Royal Society of London, A400, 1992, pp. 73-90.
- [Factor] P. Shor, *Algorithms for quantum computation: discrete logarithms and factoring*, Proceedings, 35th Annual Symposium on Fundamentals of Computer Science (FOCS), 1994, pp. 124-134.
- [Gensrch] L. Grover, *A framework for fast quantum mechanical algorithms*, lanl e-print quant-ph/9711043, to be presented in ACM Symposium on Theory of Computation, May 1998 (STOC '98).
- [Gensrch] L. Grover, *Quantum computers can search rapidly by using almost any transformation*, lanl e-print quant-ph/9712011.
- [Hogg] T. Hogg, *A framework for structured quantum search*, lanl e-print quant-ph/9701013.
- [Phone] G. Brassard, *Searching a quantum phone book*, Science, Jan. 31, 1997, 627-628.
- [Revers] C. H. Bennett, *Space-Time trade-offs in reversible computing*, SIAM Journal of Computing, vol. 18, pages 766-776, 1989.

- [Search] L. K. Grover, *Quantum Mechanics helps in searching for a needle in a haystack*, Phys. Rev. Letters, vol. 78(2), 1997, pages 325-328 (lanl e-print quant-ph/9605043).
- [Search] L. K. Grover, *A fast quantum mechanical algorithm for database search*, Proceedings 28th Annual Symposium on the Theory of Computing (STOC) 1996, pp. 212-219.
- [Structure] E. Farhi & S. Gutmann, *Quantum mechanical square root speedup in a structured search problem*, lanl e-print quant-ph/9711035.

Generalized Grover Search Algorithm for Arbitrary Initial Amplitude Distribution

David Biron¹, Ofer Biham¹, Eli Biham², Markus Grassl³, and Daniel A. Lidar⁴

¹ Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

² Computer Science Department, Technion, Haifa 32000, Israel

³ Institut für Algorithmen und Kognitive Systeme, Universität Karlsruhe,
Am Fasanengarten 5, D-76128 Karlsruhe, Germany

⁴ Department of Chemistry, University of California, Berkeley, CA 94720, USA

Abstract. Grover's algorithm for quantum searching of a database is generalized to deal with arbitrary initial amplitude distributions. First order linear difference equations are found for the time evolution of the amplitudes of the r marked and $N-r$ unmarked states. These equations are solved *exactly*. An expression for the optimal measurement time $T \sim O(\sqrt{N/r})$ is derived which is shown to depend only on the initial average amplitudes of the marked and unmarked states. A bound on the probability of measuring a marked state is derived, which depends only on the standard deviation of the initial amplitude distributions of the marked or unmarked states.

Keywords: Quantum searching, Grover's algorithm, exact solution.

1 Introduction

The power of Quantum Computation (QC) was most dramatically demonstrated in the algorithms of Shor, for the polynomial time solution of the factorization problem [1], and of Grover [2,3], for a search which can find a marked element in an unsorted database of size N , in $O(\sqrt{N})$ steps (compared to $O(N)$ steps on a classical computer). The importance of Grover's result stems from the fact that it proves the enhanced power of quantum computers compared to classical ones for a whole class of problems, for which the bound on the efficiency of classical algorithms is known. This is unlike the case of Shor's algorithm, since in spite of the fact that no efficient classical algorithm for the factorization problem is known, there is no proof that such an algorithm does not exist.

A large number of related results followed Grover's original paper [2]. Among these, the efficiency of Grover's algorithm was analyzed and compared to the theoretical efficiency limit of quantum computers for such benchmark search problems as introduced (before Grover's result [2]) by Bennett et al. [4]. The algorithm was recently shown to be optimal, i.e., to satisfy the theoretical limit [5]. Further developments include the use of Grover's algorithm or slightly modified versions of it as the essential step in algorithms that solve a variety of other problems such as quantum search for the median [6] and the minimum [7] in a set of N items, as well as the collision problem [8]. It was also shown that other

search problems which classically require $\log_2 N$ evaluations (*queries*) of a black-box function, can be reduced to a single query using Grover's approach [9,10]. Finally, it was shown that a simple closed formula describes the time evolution of the amplitudes of the generalized problem, which includes several marked states [11]. As this work is directly relevant to ours, we briefly summarize some of its pertinent results.

Let $k(t)$ [$l(t)$] denote the amplitude of the *marked* [*unmarked*] states in the database, r the number of marked states, and $\omega = 2\arcsin(\sqrt{r/N})$. It was shown by Boyer et al. [11] that after t steps of the algorithm, the marked states' amplitude increases as: $k(t) = \sin[\omega(t + 1/2)]/\sqrt{r}$, while at the same time that of the unmarked states decreases as: $l(t) = \cos[\omega(t + 1/2)]/\sqrt{N - r}$. Since N is large, the optimal time to measure and complete the calculation is thus after $T = O(\sqrt{N/r})$ time steps, when $k(t)$ is maximal. This analysis relies on the fact that the initial amplitude distribution is *uniform*. However, in a variety of interesting cases it would be desirable to apply Grover's algorithm to a *non-uniform* initial distribution. Generically, this could arise in situations where the search is used as a subroutine in a larger quantum computation, and the input to the algorithm can thus not be controlled. Another example would be cases where the given initial distribution over the elements is intrinsically non-uniform.

In this paper we generalize Grover's algorithm to the case in which the initial amplitudes are either real or complex and follow *any arbitrary distribution*. The time evolution of the amplitudes is solved *exactly* for general initial conditions, and the efficiency of the algorithm is evaluated. It is found that for generic initial conditions, the search algorithm still requires $O(\sqrt{N/r})$ steps, with only a constant factor compared to the case of a uniform initial distribution [3].

The paper is organized as follows. In Sec. 2 we define the modified Grover algorithm and derive difference equations for the time evolution of the amplitudes in it. We solve these equations exactly in Sec. 3 and analyze the results in Sec. 4. A summary and conclusions are presented in Sec. 5.

2 The Recursion Equations

2.1 The Generalized Algorithm

Our modified algorithm is essentially Grover's original algorithm, but without the initialization step. It thus consists of the following stages:

1. Use any initial distribution of marked and unmarked states, e.g., the final state of any other quantum algorithm (do *not* initialize the system to the uniform distribution).
2. Repeat the following steps T times [an expression for T is given in Eq. (24)]:
 - A. Rotate the marked states by a phase of π radians.
 - B. Rotate all states by π radians around the average amplitude of *all* states. This is done by applying the "inversion about average" operator, represented by the following unitary matrix:

$$D_{i,j} = \begin{cases} \frac{2}{N} & \text{if } i \neq j \\ \frac{2}{N} - 1 & \text{if } i = j \end{cases}$$

3. Measure the resulting state.

2.2 The Dynamics

We will now analyze the time evolution of the amplitudes in the modified algorithm with a total of N states. Let the marked amplitudes at time t be denoted by $k_i(t)$, $i = 1, \dots, r$ and the unmarked amplitudes by $l_i(t)$, $i = r + 1, \dots, N$, where the initial distribution at $t = 0$ is arbitrary. Without loss of generality we assume that the number of marked states satisfies $1 \leq r \leq N/2$. We denote the averages and variances of the amplitudes by:

$$\text{marked: } \bar{k}(t) = \frac{1}{r} \sum_{j=1}^r k_j(t) \quad \sigma_k^2(t) = \frac{1}{r} \sum_{j=1}^r |k_j(t) - \bar{k}(t)|^2 \quad (1)$$

$$\text{unmarked: } \bar{l}(t) = \frac{1}{N-r} \sum_{j=r+1}^N l_j(t) \quad \sigma_l^2(t) = \frac{1}{N-r} \sum_{j=r+1}^N |l_j(t) - \bar{l}(t)|^2 \quad (2)$$

The key observation is that the entire dynamics dictated by Grover's algorithm can be described in full by the time-dependence of the *averages*. (The variances are defined above for convenience, as they are used later in a different context – see Section 4.3.) Formally, let:

$$C(t) = -\frac{2}{N} \left[\sum_{j=1}^r k_j(t) - \sum_{j=r+1}^N l_j(t) \right] = \frac{2}{N} [(N-r)\bar{l}(t) - r\bar{k}(t)] . \quad (3)$$

$C(t)$ is thus the weighted average over the marked and unmarked states, with the minus sign accounting for the π phase difference between them during the algorithm iterations. The following theorem then shows that all states evolve equally:

Theorem 1. *The time evolution of all amplitudes (of both marked and unmarked states) is independent of the state index, and satisfies:*

$$k_i(t+1) = C(t) + k_i(t) \quad i = 1, \dots, r \quad (4)$$

$$l_i(t+1) = C(t) - l_i(t) \quad i = r+1, \dots, N \quad (5)$$

Proof. – This follows directly from the algorithm. Consider any marked state $k_i(t)$; this state is flipped to $k'_i(t) = -k_i(t)$, so that the marked average becomes $\bar{k}'(t) = \frac{1}{r} \sum_{j=1}^r k'_j(t) = -\bar{k}(t)$. The unmarked states, on the other hand, do not flip, so that the total average after the flip is: $x(t) = \frac{1}{N} [r \bar{k}'(t) + (N-r) \bar{l}(t)] = C(t)/2$. “Inversion about average” is by definition: $k'_i(t) \rightarrow 2x(t) - k'_i(t)$ and $l_i(t) \rightarrow 2x(t) - l_i(t)$. Therefore in total: $k_i(t) \rightarrow C(t) + k_i(t)$ and $l_i(t) \rightarrow C(t) - l_i(t)$. \square

From this it follows by averaging that:

Corollary 1. *The average marked and unmarked amplitudes obey first order linear coupled difference equations:*

$$\bar{k}(t+1) = C(t) + \bar{k}(t) \quad (6)$$

$$\bar{l}(t+1) = C(t) - \bar{l}(t). \quad (7)$$

These equations can be solved for $\bar{k}(t)$ and $\bar{l}(t)$, and along with the initial distribution this yield the exact solution for the dynamics of all amplitudes by using Eqs. (4) and (5).

3 Solution of the Recursion Equations

The recursion formulae can be solved by a standard diagonalization method for arbitrary complex initial conditions. Let:

$$\mathbf{v}(t) = (\bar{k}(t), \bar{l}(t)),$$

and define:

$$a \equiv \frac{N-2r}{N}, \quad b \equiv \frac{2(N-r)}{N}, \quad c \equiv \frac{2r}{N}.$$

The recursion equations (6) and (7) can be written as:

$$\mathbf{v}(t+1) = \mathbf{A} \cdot \mathbf{v}(t), \quad \mathbf{A} = \begin{pmatrix} a & b \\ -c & a \end{pmatrix}.$$

Diagonalization of \mathbf{A} yields a solution for $\mathbf{v}(t)$, as follows. Let \mathbf{S} be the diagonalizing matrix:

$$\mathbf{A}^D \equiv \mathbf{S}^{-1} \mathbf{A} \mathbf{S} = \begin{pmatrix} \lambda_- & 0 \\ 0 & \lambda_+ \end{pmatrix}, \quad \lambda_{\pm} = \gamma e^{\pm i\omega}.$$

Then $\mathbf{w}(t) = \mathbf{S}^{-1} \cdot \mathbf{v}(t)$ satisfies:

$$\mathbf{w}(t+1) = \mathbf{A}^D \cdot \mathbf{w}(t),$$

with solution:

$$\mathbf{w}(t) = ((\lambda_-)^t w_-(0), (\lambda_+)^t w_+(0))$$

where $\mathbf{w}(0) = (w_-(0), w_+(0))$. This yields $\bar{k}(t)$ and $\bar{l}(t)$ from $\mathbf{v}(t) = \mathbf{S} \cdot \mathbf{w}(t)$. Diagonalizing \mathbf{A} one finds:

$$\gamma = a^2 + bc = 1 \quad (8)$$

$$\cos \omega = a = 1 - 2\frac{r}{N}, \quad (9)$$

which is identical to the frequency found by Boyer et al. [11] The eigenvectors of \mathbf{A} are the columns of \mathbf{S} :

$$\mathbf{S} = \begin{pmatrix} i\sqrt{\frac{N}{r}-1} & -i\sqrt{\frac{N}{r}-1} \\ 1 & 1 \end{pmatrix}, \quad \mathbf{S}^{-1} = \begin{pmatrix} -\frac{i}{2}\sqrt{\frac{r}{N-r}} & \frac{1}{2} \\ \frac{i}{2}\sqrt{\frac{r}{N-r}} & \frac{1}{2} \end{pmatrix}.$$

Using this:

$$\begin{pmatrix} w_{-}(0) \\ w_{+}(0) \end{pmatrix} = \mathbf{w}(0) = \mathbf{S}^{-1} \cdot \mathbf{v}(0) = \begin{pmatrix} -\frac{i}{2}\sqrt{\frac{r}{N-r}}\bar{k}(0) + \frac{1}{2}\bar{l}(0) \\ \frac{i}{2}\sqrt{\frac{r}{N-r}}\bar{k}(0) + \frac{1}{2}\bar{l}(0) \end{pmatrix},$$

so that:

$$\mathbf{v}(t) = \mathbf{S} \cdot \begin{pmatrix} \left(-\frac{i}{2}\sqrt{\frac{r}{N-r}}\bar{k}(0) + \frac{1}{2}\bar{l}(0)\right) e^{-i\omega t} \\ \left(\frac{i}{2}\sqrt{\frac{r}{N-r}}\bar{k}(0) + \frac{1}{2}\bar{l}(0)\right) e^{i\omega t} \end{pmatrix}.$$

This yields finally, after some straightforward algebra:

$$\bar{k}(t) = \bar{k}(0) \cos \omega t + \bar{l}(0) \sqrt{\frac{N-r}{r}} \sin \omega t \quad (10)$$

$$\bar{l}(t) = \bar{l}(0) \cos \omega t - \bar{k}(0) \sqrt{\frac{r}{N-r}} \sin \omega t. \quad (11)$$

Together with Eqs. (4) and (5) this provides the complete exact solution to the dynamics of the amplitudes in the generalized Grover algorithm, for arbitrary initial conditions.

4 Analysis

Next we derive several properties of the amplitudes.

4.1 Phase Difference

The averaged amplitudes can be expressed concisely as follows (even when $\bar{k}(0)$ and $\bar{l}(0)$ are complex):

$$\bar{k}(t) = \alpha \sin(\omega t + \phi) \quad (12)$$

$$\bar{l}(t) = \beta \cos(\omega t + \phi) \quad (13)$$

where

$$\begin{aligned} \tan \phi &= \frac{\bar{k}(0)}{\bar{l}(0)} \sqrt{\frac{r}{N-r}}; & \alpha^2 &= \bar{k}(0)^2 + \bar{l}(0)^2 \frac{N-r}{r}; \\ & & \beta^2 &= \bar{l}(0)^2 + \bar{k}(0)^2 \frac{r}{N-r} \end{aligned} \quad (14)$$

which shows that there is a $\pi/2$ phase difference between the marked and unmarked amplitudes: when the average marked amplitude is maximal, the average unmarked amplitude is minimal, and *vice versa*.

4.2 Constant Variance

Subtracting Eq. (4) from Eq. (6), and subtracting Eq. (5) from Eq. (7), one finds:

$$k_i(t+1) - \bar{k}(t+1) = k_i(t) - \bar{k}(t) \quad (15)$$

$$l_i(t+1) - \bar{l}(t+1) = -[l_i(t) - \bar{l}(t)]. \quad (16)$$

This means that:

$$\Delta k_i \equiv k_i(t) - \bar{k}(t) \quad \text{and} \quad \Delta l_i \equiv (-1)^t [l_i(t) - \bar{l}(t)], \quad (17)$$

are *constants of motion* (time-independent). It follows immediately from the definition that the variances σ_k^2 and σ_l^2 (cf. Eqs. (eq:marked) and (eq:unmarked)) too, are both time-independent.

This allows us to simplify the expression for the time dependence of the amplitudes:

$$k_i(t) = \bar{k}(t) + \Delta k_i \quad (18)$$

$$l_i(t) = \bar{l}(t) + (-1)^t \Delta l_i, \quad (19)$$

where Δk_i and Δl_i are evaluated at $t = 0$.

4.3 Maximal Probability of Success and Optimal Number of Iterations

The probability that a marked state will be obtained in the measurement at time t at the end of the process is $P(t) = \sum_{j=1}^r |k_j(t)|^2$. A bound on this quantity can be derived as follows. Since all the operators used are unitary, the amplitudes satisfy the normalization condition:

$$\sum_{i=1}^r |k_i(t)|^2 + \sum_{i=r+1}^N |l_i(t)|^2 = 1 \quad (20)$$

at all times. Using $\overline{(y - \bar{y})^2} = \bar{y}^2 - \bar{y}^2$ (y is a random variable), we find from Eq. (2):

$$\sum_{i=r+1}^N |l_i(t)|^2 = (N-r)\sigma_l^2 + \left| \sum_{i=r+1}^N l_i(t) \right|^2 / (N-r).$$

Let:

$$P_{\max} = 1 - (N-r)\sigma_l^2, \quad (21)$$

a time-independent quantity. Note that in the case of uniform initial distribution of amplitudes $\sigma_l^2 = 0$ and $P_{\max} = 1$. Now, $P(t) = P_{\max} - (N - r)|\bar{l}(t)|^2$, so that:

$$P(t) \leq P_{\max} \quad (22)$$

is the required bound. Using the exact solution, we can show that the P_{\max} bound is in fact tight. For, from Eq. (13) it follows that $\bar{l}(T) = 0$ when:

$$\omega T + \phi = (j + 1/2)\pi, \quad j = 0, 1, 2, \dots \quad (23)$$

At these times the bound is reached so that times T satisfying Eq. (23) are optimal for measurement. Note that this conclusion holds only if $\bar{k}(0)/\bar{l}(0)$ is real. When $\bar{k}(0)/\bar{l}(0)$ is complex, the bound is generally not reached since $\bar{l}(t)$ may never vanish. Collecting our results:

Theorem 2. *Given arbitrary initial distributions of r marked and $N - r$ unmarked states, with known averages $\bar{k}(0)$ and $\bar{l}(0)$ respectively, $\bar{k}(0)/\bar{l}(0)$ real, the optimal measurement times are after:*

$$T = \frac{(j + 1/2)\pi - \arctan \left[\frac{\bar{k}(0)}{\bar{l}(0)} \sqrt{\frac{r}{N-r}} \right]}{\arccos \left(1 - 2\frac{r}{N} \right)}, \quad j = 0, 1, 2, \dots \quad (24)$$

steps, when the probability of obtaining a marked state is P_{\max} as given by Eq. (21).

An important conclusion is that to determine the optimal measurement times, all one needs to know are the average initial amplitudes and the number of marked states. The more difficult case when these are unavailable will be considered in a separate publication [12]. The expansion of Eq. (24) in $r/N \ll 1$ (at $j = 0$) yields:

$$T = -\frac{1}{2} \frac{\bar{k}(0)}{\bar{l}(0)} + \frac{\pi}{4} \sqrt{N/r} - \frac{\pi}{24} \sqrt{r/N} + O(r/N), \quad (25)$$

confirming that Grover's algorithm converges in $O(\sqrt{N/r})$ steps for arbitrary distributions. The advantage of an initial amplitude distribution with a relatively high average of the marked states is manifested in the constant offset $-\frac{1}{2} \frac{\bar{k}(0)}{\bar{l}(0)}$, which may significantly reduce the required number of steps.

5 Summary and Conclusions

In this work we generalized Grover's quantum search algorithm to apply for initial input distributions which are non-uniform. In fact, it was shown that by simply omitting the first step of Grover's original algorithm, wherein a uniform superposition is created over all elements in the database, a more general algorithm results which applies to *arbitrary* initial distributions. To analyze the algorithm, we found that the time evolution of the amplitudes of the marked

and unmarked states can be described by first-order linear difference equations with some special properties. The most important of these is that all amplitudes essentially evolve uniformly, with the dynamics being determined completely by the average amplitudes. This observation allowed us to find an exact solution for the time-evolution of the amplitudes. An important conclusion from this solution is that generically the generalized algorithm also has a $O(\sqrt{N/r})$ running time, thus being more powerful than any classical algorithm designed to solve the same task.

This work was initiated during the 1997 Elsag-Bailey – I.S.I. Foundation research meeting on quantum computation.

References

1. P.W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM Journal on Computing*, **26**, 1484 (1997).
2. L. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-Eighth Annual Symposium on the Theory of Computing*, ACM Press (New York, 1996), p. 212.
3. L. Grover, Quantum mechanics helps in searching for a needle in a haystack, *Phys. Rev. Lett.* **79**, 325 (1997).
4. C.H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, Strengths and weaknesses of quantum computing, *SIAM Journal on Computing* **26**, 1510 (1997).
5. C. Zalka, Grover's quantum searching algorithm is optimal (LANL preprint quant-ph/9711070).
6. L. Grover, Quantum telecomputation (LANL preprint quant-ph/9704012).
7. C. Durr and P. Hoyer, A quantum algorithm for finding the minimum (LANL preprint quant-ph/9607014).
8. G. Brassard, P. Hoyer, and A. Tapp, Quantum algorithm for the collision problem (LANL preprint quant-ph/9705002).
9. L. Grover, Quantum computers can search arbitrarily large databases by a single query, *Phys. Rev. Lett.* **79**, 4709 (1997).
10. B.M. Terhal and J.A. Smolin, Single quantum querying of a database (LANL preprint quant-ph/9705041).
11. M. Boyer, G. Brassard, P. Hoyer and A. Tapp, Tight bounds on quantum searching, in *Proceedings of the fourth workshop on Physics and Computation*, edited by T. Toffoli, M. Biafore and J. Leao, New England Complex Systems Institute, (Boston, 1996), p. 36. To appear in *Fortschritte der Physik*.
12. D. Biron, O. Biham, E. Biham, M. Grassl, and D.A. Lidar, to be published.

Quantum Database Search by a Single Query

Dong Pyo Chi and Jinsoo Kim

Department of Mathematics, Seoul National University,
Seoul 151-742, Korea.

Abstract. In this paper we give a quantum mechanical algorithm that can search a database by a single query when the number of solutions is more than a quarter. It utilizes the generalized Grover operator of arbitrary phase.

1 Introduction

For $N \in \mathbb{N}$ let $\mathbb{Z}_N = \{0, 1, \dots, N-1\}$ denote the additive cyclic group of order N and let $F : \mathbb{Z}_N \rightarrow \mathbb{Z}_2$ be the Boolean function computed by the oracle. The database search problem is to find some $j \in \mathbb{Z}_N$ such that $F(j) = 1$ under the assumption that such an j exists. We assume that the structure of F is unknown so that it is not possible to obtain a knowledge about F without evaluating it on \mathbb{Z}_N .

Grover [4] constructed a quantum mechanical algorithm that can solve the database search problem in expected time of order $O(\sqrt{N/t})$ where $t = |\{j \in \mathbb{Z}_N : F(j) = 1\}|$ [2]. Bennett *et al.* [1] have shown that the work of Grover is optimal up to a multiplicative constant. Especially when $t = N/4$ is known, the original Grover algorithm in [4] can search a solution only by a single query [2]. It uses the π -phase, i.e., marking the states by multiplying $e^{\pi i} = -1$. When $t = N/2$, by changing this phase to $\pi/2$, that is, by marking the states by multiplying $e^{\pi i/2} = i$ and modifying the corresponding diffusion transform according to this phase, the solution can be found with certainty after a single iteration [3].

In this paper we generalize Grover algorithm in order to permit arbitrary phase. When $t \geq N/4$, we give generalized conditional phase and diffusion transform depending on t and then formulate a quantum mechanical algorithm that solves the database search problem in a single query.

2 Generalized Grover Operator

Let $\mathcal{B}_N = \{|a\rangle\}_{a \in \mathbb{Z}_N}$ be the standard basis of an n -qubit quantum register with $N = 2^n$ and \mathcal{H}_N be the corresponding Hilbert space which represents the state vectors of the quantum system. Let l be any element in \mathbb{Z}_N .

For $\gamma \in \mathbb{R}$ the *conditional γ -phase transform* $\mathbf{S}_{F,\gamma} : \mathcal{H}_N \rightarrow \mathcal{H}_N$ is defined by

$$\mathbf{S}_{F,\gamma}|j\rangle = (e^{i\gamma})^{F(j)}|j\rangle$$

for $j = 0, 1, \dots, N-1$. Let $\mathbf{S}_{l,\gamma}$ denote $\mathbf{S}_{F_l,\gamma}$ where $F_l(j) = \delta_{jl}$.

Let \mathbf{W}_l be any unitary transformation on \mathcal{H}_N satisfying

$$\mathbf{W}_l|l\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle.$$

For example, we may set \mathbf{W}_0 to be the Walsh-Hadamard transform. When N is not a power of 2 the approximate Fourier transform in [5] can be used.

For $\beta \in \mathbb{R}$ the β -phase diffusion transform $\mathbf{D}_\beta : \mathcal{H}_N \rightarrow \mathcal{H}_N$ is defined by

$$\begin{aligned} \mathbf{D}_\beta &= \mathbf{W}_l \mathbf{S}_{l,\beta} \mathbf{W}_l^\dagger \\ &= \mathbf{I} + (e^{i\beta} - 1)\mathbf{P}, \end{aligned}$$

where \mathbf{I} is the identity matrix and $\mathbf{P} = (P_{jk})$ is a projection matrix in \mathcal{H}_N with $P_{jk} = \frac{1}{N}$. Notice that the definition of \mathbf{D}_β does not depend on l . We note that $\mathbf{S}_{F,\gamma}$ and \mathbf{D}_β are unitary.

Let us define the *generalized Grover operator of (β, γ) -phase* $\mathbf{G}_{F,\beta,\gamma} : \mathcal{H}_N \rightarrow \mathcal{H}_N$ by

$$\mathbf{G}_{F,\beta,\gamma} = \mathbf{D}_\beta \mathbf{S}_{F,\gamma}.$$

When $\beta = \gamma$ we set $\mathbf{G}_{F,\gamma} = \mathbf{G}_{F,\gamma,\gamma}$.

Let

$$A = \{j \in \mathbb{Z}_N | F(j) = 1\} \quad \text{and} \quad B = \{j \in \mathbb{Z}_N | F(j) = 0\}.$$

Then $t = |A|$. For $k_0, l_0 \in \mathbb{C}$ such that $t|k_0|^2 + (N-t)|l_0|^2 = 1$ let us define

$$|\psi(k_0, l_0)\rangle = \sum_{j \in A} k_0 |j\rangle + \sum_{j \in B} l_0 |j\rangle$$

and let k_j and l_j be the corresponding amplitudes after j iterations of the generalized Grover operator of (β, γ) -phase;

$$|\psi(k_j, l_j)\rangle = \mathbf{G}_{F,\beta,\gamma}^j |\psi(k_0, l_0)\rangle.$$

Then we have the following lemma.

Lemma 1. *For a nonnegative integer j after applying $j + 1$ the generalized Grover operator of (β, γ) -phase to the initial state $|\psi(k_0, l_0)\rangle$, the state becomes $|\psi(k_{j+1}, l_{j+1})\rangle$ where*

$$\begin{cases} k_{j+1} = \frac{(e^{i\beta} - 1)t + N}{N} e^{i\gamma} k_j + \frac{(e^{i\beta} - 1)(N - t)}{N} l_j, \\ l_{j+1} = \frac{(e^{i\beta} - 1)t}{N} e^{i\gamma} k_j + \frac{(e^{i\beta} - 1)(N - t) + N}{N} l_j. \end{cases} \quad (1)$$

3 Quantum Database Search

Theorem 1. Assume that $k_0 = l_0$ and $\beta, \gamma \in [0, 2\pi]$. Then $l_1 = 0$ if and only if $\frac{N}{4} \leq t \leq N$ and $\beta = \gamma = \cos^{-1}(1 - \frac{N}{2t})$. In this case, we have $k_1 = (e^{i\gamma} - 1)k_0$ and $\beta, \gamma \in [\pi/3, 5\pi/3]$.

Proof. By (1) we get

$$\begin{cases} k_1 = \left[(e^{i\beta} - 1)(e^{i\gamma} - 1) \frac{t}{N} + e^{i\gamma} + e^{i\beta} - 1 \right] k_0, \\ l_1 = \left[(e^{i\beta} - 1)(e^{i\gamma} - 1) \frac{t}{N} + e^{i\beta} \right] l_0. \end{cases} \quad (2)$$

Considering the imaginary part of $e^{-i\beta}l_1/l_0$, the equation

$$\frac{t}{N} \{ (1 - \cos \beta) \sin \gamma + (\cos \gamma - 1) \sin \beta \} = 0$$

is equivalent to

$$\frac{1 - \cos \beta}{\sin \beta} = \frac{1 - \cos \gamma}{\sin \gamma}. \quad (3)$$

Considering the real part of $e^{-i\beta}l_1/l_0$, it follows from (3) that the equation

$$\frac{t}{N} \{ (1 - \cos \beta)(1 - \cos \gamma) + \sin \beta \sin \gamma \} - 1 = 0$$

is equivalent to

$$\cos \beta = \cos \gamma = 1 - \frac{N}{2t}.$$

Thus we get $\beta, \gamma \in [\pi/3, 5\pi/3]$, $\beta = \gamma$ and $t \geq \frac{N}{4}$ by (3) again. Furthermore, by (2) we obtain $k_1 = (e^{i\gamma} - 1)k_0$. This completes the proof. \square

For the case of π -phase in [4], $-\mathbf{D}_\pi = -\mathbf{I} + 2\mathbf{P}$ is an inversion about average operation and we have

$$-\mathbf{D}_\pi |\psi(k, l)\rangle = \left| \psi \left(-\frac{N-2t}{N}k + \frac{2(N-t)}{N}l, \frac{N-2t}{N}l + \frac{2t}{N}k \right) \right\rangle.$$

In this case, there is an explicit closed-form formula for k_j and l_j [2];

$$\begin{cases} k_j = (-1)^j \frac{1}{\sqrt{t}} \sin((2j+1)\theta), \\ l_j = (-1)^j \frac{1}{\sqrt{N-t}} \cos((2j+1)\theta) \end{cases}$$

for all nonnegative integer j where the angle θ is defined so that $\sin^2 \theta = \frac{t}{N}$. Especially when $t = N/4$ we have $l_1 = 0$.

Grover operator of $\frac{\pi}{2}$ -phase was used in [3]. Since

$$\mathbf{D}_{\frac{\pi}{2}}|\psi(k, l)\rangle = \left| \psi \left(\frac{(i-1)t+N}{N}k + \frac{(i-1)(N-t)}{N}l, \frac{(i-1)(N-t)+N}{N}l + \frac{(i-1)t}{N}k \right) \right\rangle,$$

when $t = \frac{N}{2}$ we have

$$\mathbf{G}_{F, \frac{\pi}{2}}|\psi(k, k)\rangle = |\psi((i-1)k, 0)\rangle.$$

When $t \in [N/4, N]$ by Theorem 1 we have

$$\mathbf{G}_{F, \gamma}|\psi(k_0, k_0)\rangle = |\psi((e^{i\gamma} - 1)k_0, 0)\rangle$$

where the phase γ is defined by $\gamma = \cos^{-1} \left(1 - \frac{N}{2t} \right)$. Thus a single iteration of the generalized Grover operator makes all the amplitudes of the bases in B vanish and we can find a solution by a single query.

References

1. Bennett, C. H., Bernstein, E., Brassard, G., Vazirani, U.: Strengths and weaknesses of quantum computing. *SIAM J. Comput.* **26** (1997) 1510–1523
2. Boyer, M., Brassard, G., Høyer, P., Tapp, A.: Tight bounds on quantum searching. *Proceedings of the Fourth Workshop on Physics and Computation*, New England Complex Systems Institute, 1996, 36–43
3. Brassard, G., Høyer, P.: An exact quantum polynomial-time algorithm for Simon's problem. *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems*, 1997 (to appear)
4. Grover, L. K.: Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79** (1997) 325–328
5. Kitaev, A. Yu.: Quantum measurements and the abelian stabilizer problem. Los Alamos e-print quant-ph/9511026, 1995

Quantum Computer Can Not Speed Up Iterated Applications of a Black Box

Y. Ozhigov

Department of mathematics, Moscow state technological University "Stankin",
Vadkovsky per. 3a, 101472, Moscow, Russia, y@oz.msk.ru

Abstract. Let a classical algorithm be determined by sequential applications of a black box performing one step of this algorithm. If we consider this black box as an oracle which gives a value $f(a)$ for a query a , we can compute T sequential applications of f on a classical computer relative to this oracle in time T .

It is proved that if $T = O(2^{n/7})$, where n is the length of input, then the result of T sequential applications of f can not be computed on quantum computer with oracle for f for all possible f faster than in time $\Omega(T)$. This means that there is no general method of quantum speeding up of classical algorithms provided in such a general method a classical algorithm is regarded as iterated applications of a given black box.

1 Introduction

In the last years many investigators have amassed a convincing body of evidence that a quantum device can be more powerful tool for computations than a classical computer. This is because for the different problems there exist quantum algorithms which find a solution substantially faster than any known (or even any possible) classical algorithm (look, for example, at the works [DJ], [BB], [Sh]). The latest advance in quantum speeding up is the method of quantum search proposed by L.Grover in the work [Gr]. His algorithm takes $O(\sqrt{N})$ time when the classical search requires $\Omega(N)$ time. In some particular cases (look in [FG]) the time $O(\sqrt{N})$ for a search can be even reduced. It would be natural to expect that some more general method of quantum speeding up can take place for all classical algorithms with time complexity more than $O(n)$.

One of the main general corollaries from the classical theory of algorithms is that if we know only a code of algorithm then in general case the unique way to learn a result of computations is to run this algorithm on a given input. Therefore, given a code of algorithm, generally speaking we can only use it as a black box to perform sequentially all steps of computations and no other analysis can yield their result. Thus we can regard a computation $X_0 \rightarrow X_1 \rightarrow \dots \rightarrow X_T$ as iterated application of the same oracle f which gives sequentially $X_{s+1} = f(X_s)$, $s = 0, 1, \dots, T-1$, $T = T(n) > O(n)$.

In view of this we assume that a general method of quantum speeding up of classical algorithms is a quantum query machine with oracle f which yields

the result X_T of computations in time $\alpha(T)$, where $\alpha(T)/T \rightarrow 0$ ($T \rightarrow \infty$). However, we shall see that such a method does not exist. This demonstrates a value of every partial result about quantum speeding up because such results are all that can be done.

Oracle quantum computers will be treated here within the framework of approach proposed by C.Bennett, E.Bernstein, G.Brassard and U.Vazirani in the work [BBBV]. They considered a quantum Turing machine with oracle as a model of quantum computer. In this paper we use slightly different model of quantum computer with separated quantum and classical parts, but the results hold also for QTMs. We proceed with the exact definitions.

2 Quantum Computer with the Separated Quantum and Classical Parts

Our quantum query machine consists of two parts: quantum and classical. Let ω^* denotes the set of all words in alphabet ω .

Quantum part.

It consists of two infinite tapes: working and query, the finite set \mathcal{U} of unitary transformations which can be easily performed by the physical devices, and infinite set $F = \bigcup_{n=1}^{\infty} F_n$ of unitary transformations called an oracle for the length preserving function $f : \{0,1\}^* \rightarrow \{0,1\}^*$, each F_n acts on 2^{2n} dimensional Hilbert space spanned by $\{0,1\}^{2n}$ as follows: $F_n|\bar{a}, \bar{b}\rangle = |\bar{a}, f(\bar{a}) \oplus \bar{b}\rangle$, $\bar{a}, \bar{b} \in \{0,1\}^n$, where \oplus denotes the bitwise addition modulo 2.

The cells of tapes are called qubits. Each qubit takes values from the complex 1-dimensional sphere of radius 1: $\{z_0\mathbf{0} + z_1\mathbf{1} \mid z_1, z_2 \in \mathbb{C}, |z_0|^2 + |z_1|^2 = 1\}$. Here $\mathbf{0}$ and $\mathbf{1}$ are referred as basic states of qubit and form the basis of \mathbb{C}^2 .

During all the time of computation the both tapes are limited each by two markers with fixed positions, so that on the working (query) tape only qubits v_1, v_2, \dots, v_T ($v_{T+1}, v_{T+2}, \dots, v_{T+2n}$) are available in a computation with time complexity $T = T(n)$ on input of length n . Put $Q = \{v_1, v_2, \dots, v_{T+2n}\}$. A basic state of quantum part is a function of the form $e : Q \rightarrow \{0,1\}$. Such a state can be encoded as $|e(v_1), e(v_2), \dots, e(v_{T+2n})\rangle$ and naturally identified with the corresponding word in alphabet $\{0,1\}$. Let $K = 2^{T+2n}$; e_0, e_1, \dots, e_{K-1} be all basic states taken in some fixed order, \mathcal{H} be K dimensional Hilbert space with orthonormal basis e_0, e_1, \dots, e_{K-1} . \mathcal{H} can be regarded as tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_{T+2n}$ of 2 dimensional spaces, where \mathcal{H}_i is generated by all possible values of v_i , $i = 1, 2, \dots, T+2n$. A (pure) state of quantum part is such an element $x \in \mathcal{H}$ that $|x| = 1$.

Time evolution of quantum part at hand is determined by two types of unitary transformations on its states: working and query. Let a pair G, U be somehow selected, where $G \subset \{1, 2, \dots, T+2n\}$, $U \in \mathcal{U}$ is unitary transform on $2^{\text{card}(G)}$ dimensional Hilbert space.

Working transform $W_{G,U}$ on \mathcal{H} has the form $E \otimes U'$, where U' acts as U on $\bigotimes_{i \in G} \mathcal{H}_i$ in the basis at hand, E acts as identity on $\bigotimes_{i \notin G} \mathcal{H}_i$.

Query transform Qu_f on \mathcal{H} has the form $E \otimes F'_n$, where F'_n acts as F_n on $\bigotimes_{i=T+1}^{T+2n} \mathcal{H}_i$ and E acts as identity on $\bigotimes_{i=1}^T \mathcal{H}_i$.

Observation of the quantum part. If the quantum part is in state $\chi = \sum_{i=0}^{K-1} \lambda_i e_i$, an observation is a procedure which gives the basic state e_i with probability $|\lambda_i|^2$.

Classical part.

It consists of two classical tapes: working and query, which cells are in one-to-one correspondence with the respective qubits of the quantum tapes and have boundary markers on the corresponding positions. Every cell of classical tapes contains a letter from some finite alphabet ω . Evolution of classical part is determined by the classical Turing machine M with a few heads on both tapes and the set of integrated states of heads: $\{q_b, q_w, q_q, q_o, \dots\}$. We denote by $h(C)$ the integrated state of heads for a state C of classical part.

Let D be the set of all states of classical part.

Rule of correspondence between quantum and classical parts has the form $R: D \longrightarrow 2^{\{1,2,\dots,T+2n\}} \times \mathcal{U}$, where $\forall C \in D$ $R(C) = \langle G, U \rangle$, U acts on $2^{\text{card}(G)}$ dimensional Hilbert space so that U depends only on $h(C)$, and the elements of G are exactly the numbers of those cells on classical tape which contain the special letter $a_0 \in \omega$.

A state of quantum computer at hand is a pair $S = \langle Q(S), C(S) \rangle$ where $Q(S)$ and $C(S)$ are the states of quantum and classical parts respectively.

Computation on quantum computer. It is a chain of transformations of the following form:

$$S_0 \longrightarrow S_1 \longrightarrow \dots \longrightarrow S_T, \quad (1)$$

where for every $i = 0, 1, \dots, T-1$ $C(S_i) \longrightarrow C(S_{i+1})$ is transformation determined by Turing machine M , and the following properties are fulfilled:

- if $h(C(S_i)) = q_w$ then $Q(S_{i+1}) = W_{R(C(S_i))}(Q(S_i))$,
- if $h(C(S_i)) = q_q$ then $Q(S_{i+1}) = \text{Qu}_f(Q(S_i))$,
- if $h(C(S_i)) = q_b$ then $i = 0$, $Q(S_0) = e_0$, $C(S_0)$ is fixed initial state, corresponding to input word $a \in \{0, 1\}^n$,
- if $h(C(S_i)) = q_o$ then $i = T$,
- in other cases $Q(S_{i+1}) = Q(S_i)$.

We say that this quantum computer (QC) computes a function $F(a)$ with probability $p \geq 2/3$ and time complexity T if for the computation **(II)** on every input a the observation of S_T and the following routine procedure fixed beforehand give $F(a)$ with probability p . We always can reach any other value of probability $p_0 > p$ if fulfill computations repeatedly on the same input and take the prevailing result. This leads only to a linear slowdown of computation.

3 The Effect of Changes in Oracle on the Result of Quantum Computation

For a state $e_j = |s_1, s_2, \dots, s_{T+2n}\rangle$ of the quantum part we denote the word $s_{T+1}s_{T+2} \dots s_{T+n}$ by $q(e_j)$. The state S of QC is called query if $h(C(S)) = q_q$. Such a state is querying the oracle on all the words $q(e_j)$ with some amplitudes. Put $\mathcal{K} = \{0, 1, \dots, K-1\}$. Let $\xi = Q(S) = \sum_{j \in \mathcal{K}} \lambda_j e_j$. Given a word $a \in \{0, 1\}^n$ for a query state S we define:

$$\delta_a(\xi) = \sum_{j: q(e_j)=a} |\lambda_j|^2.$$

It is the probability that a state S is querying the oracle on the word a . In particular, $\sum_{a \in \{0,1\}^n} \delta_a(\xi) = 1$.

Each query state S induces the metric on the set of all oracles if for length preserving functions f, g we define a distance between them by

$$d_S(f, g) = \left(\sum_{a: f(a) \neq g(a)} \delta_a(\xi) \right)^{1/2}.$$

Lemma 1 *Let Qu_f , Qu_g be query transforms on quantum part of QC corresponding to functions f, g ; S be a query state. Then*

$$|\text{Qu}_f(S) - \text{Qu}_g(S)| \leq 2d_S(f, g).$$

Proof

Put $\mathcal{L} = \{j \in \mathcal{K} \mid f(q(e_j)) \neq g(q(e_j))\}$. We have: $|\text{Qu}_f(S) - \text{Qu}_g(S)| \leq 2 \left(\sum_{j \in \mathcal{L}} (|\lambda_j|)^2 \right)^{1/2} \leq 2d_S(f, g)$. Lemma is proved.

Now we shall consider the classical part of computer as a part of working tape. Then a state of computer will be a point in K^2 dimensional Hilbert space \mathcal{H}_1 . We denote such states by ξ, χ with indices. All transformations of classical part can be fulfilled reversibly as it is shown by C.Bennett in the work [Be]. This results in that all transformations in computation (II) will be unitary transforms in \mathcal{H}_1 . At last we can join sequential steps: $S_i \longrightarrow S_{i+1} \longrightarrow \dots \longrightarrow S_j$ where $S_i \longrightarrow S_{i+1}$, $S_j \longrightarrow S_{j+1}$ are two nearest query transforms, in one step. So the computation on our QC acquires the form

$$\chi_0 \longrightarrow \chi_1 \longrightarrow \dots \longrightarrow \chi_t, \tag{2}$$

where every passage is the query unitary transform and the following unitary transform U_i which depends only on i : $\chi_i \xrightarrow{\text{Qu}_f} \chi'_i \xrightarrow{U_i} \chi_{i+1}$. We shall denote $U_i(\text{Qu}_f(\xi))$ by $V_{i,f}(\xi)$, then $\chi_{i+1} = V_{i,f}(\chi_i)$, $i = 0, 1, \dots, t-1$. Here t is the number of query transforms (or evaluations of the function f) in the computation at hand. Put $d_a(\xi) = \sqrt{\delta_a(\xi)}$.

Lemma 2 *If $\chi_0 \longrightarrow \chi_1 \longrightarrow \dots \longrightarrow \chi_t$ is a computation with oracle for f , a function g differs from f only on one word $a \in \{0, 1\}^n$ and $\chi_0 \longrightarrow \chi'_1 \longrightarrow \dots \longrightarrow \chi'_t$ is a computation on the same QC with a new oracle for g , then*

$$|\chi_t - \chi'_t| \leq 2 \sum_{i=0}^{t-1} d_a(\chi_i).$$

Proof

Induction on t . Basis is evident. Step. In view of that $V_{t-1,g}$ is unitary, Lemma 1 and inductive hypothesis, we have

$$\begin{aligned} |\chi_t - \chi'_t| &= |V_{t-1,f}(\chi_{t-1}) - V_{t-1,g}(\chi'_{t-1})| \leq \\ &|V_{t-1,f}(\chi_{t-1}) - V_{t-1,g}(\chi_{t-1})| + |V_{t-1,g}(\chi_{t-1}) - V_{t-1,g}(\chi'_{t-1})| \leq \\ 2d_a(\chi_{t-1}) + |\chi_{t-1} - \chi'_{t-1}| &= 2d_a(\chi_{t-1}) + 2 \sum_{i=0}^{t-2} d_a(\chi_i) = 2 \sum_{i=0}^{t-1} d_a(\chi_i). \end{aligned}$$

Lemma is proved.

4 Main Result

For a length preserving function f a result of its iteration $f^{\{k\}}$ is defined by the induction on k : $f^{\{0\}}$ is identity mapping, $f^{\{k+1\}}(x) = f(f^{\{k\}}(x))$.

Theorem 1 *There is no such QC with oracle for f that for some functions $t(n), T(n) : t(n)/T(n) \longrightarrow 0$ ($n \longrightarrow \infty$), $T(n) = O(2^{n/7})$ and every f QC computes $f^{\{T(n)\}}(\bar{0})$ applying only $t(n)$ evaluations of f .*

Proof

Suppose that it is not true and some QC with oracle for f computes $f^{\{T(n)\}}(\bar{0})$ applying only $t(n)$ evaluations of f , where $t(n)/T(n) \longrightarrow 0$ ($n \longrightarrow \infty$), $T(n) = O(2^{n/7})$, and obtain a contradiction.

Let $f : \{0, 1\}^* \longrightarrow \{0, 1\}^*$ be such length preserving bijection that for every $n = 1, 2, \dots$ the orbit of the word $\bar{0} = 0^n$ contains all words from $\{0, 1\}^n$. Let an oracle for f be taken for the computation of $f^{\{T\}}(\bar{0})$ on our QC. This computation has the form (2) where $t/T \longrightarrow 0$ ($n \longrightarrow \infty$). Let n be sufficiently large so that $5t < T$.

Now we shall define the lists of the form $\langle \xi_i, f_i, \mathcal{T}_i, x_i \rangle$ where ξ_i is a state from \mathcal{H}_1 , $|\xi_i| = 1$, $f_i : \{0, 1\}^* \longrightarrow \{0, 1\}^*$ is length preserving function, $x_i \in \mathcal{T}_i \subseteq \{0, 1\}^n$ by the following induction on i .

Basis: $i = 0$. Put $\xi_0 = \chi_0$, $f_0 = f$, $x_0 = \bar{0}$, $\mathcal{T}_0 = \{0, 1\}^n$.

Step. Put

$$\begin{aligned} \xi_{i+1} &= V_{i,f_i}(\xi_i), \\ \mathcal{T}_{i+1} &= \mathcal{T}_i \cap R_i, \quad R_i = \{a \mid \delta_a(\xi_{i+1}) < \frac{1}{T^\alpha}\}, \end{aligned}$$

f_{i+1} differs from f_i at most on one word x_i where we define $x_{i+1} = f_{i+1}(x_i)$ such that for all $s = 1, 2, \dots, T$ $f_{i+1}^{\{s\}}(x_i) \in \mathcal{T}_{i+1}$.

Note that $2^n - \text{card}(\mathbf{R}_i) < T^\alpha$. Therefore we can chose x_{i+1} such that $x_{i+1} = f^{\{j\}}(\bar{0})$ where $j < (i+1)TT^\alpha$. It is possible for every $i = 1, 2, \dots, t-1$ if $\alpha \leq 5$ and n is sufficiently large, because $T = O(2^{\frac{n}{2}})$.

We introduce the following notations: $V_i = V_{i,f_i}$, $V_i^* = V_{i,f_i}$. Let the unitary operator V^i be introduced by the following induction: $V^0(x) = V_0(x)$, $V^i(x) = V_i(V^{i-1}(x))$, and the unitary operator \tilde{V}_i be defined by $\tilde{V}_0 = V_0^*$, $\tilde{V}_i(x) = V_i^*(\tilde{V}_{i-1}(x))$. Then $\xi_{i+1} = \tilde{V}_i(\xi_0)$.

Put $\xi'_0 = \xi_0$, $\xi'_{i+1} = V^i(\xi_0)$, $\partial_i = |\xi_i - \xi'_i|$, $\Delta_i = |V_i^*(\xi_i) - V_i(\xi_i)|$. It follows from the definition that f_i differs from f_t at most on the set $X_i = \{x_i, x_{i+1}, \dots, x_{t-1}\}$ where $\forall a \in X_i \delta_a(\xi_i) < \frac{1}{T^\alpha}$. Consequently, applying Lemma 1 we obtain

$$\Delta_i \leq \frac{2t^{1/2}}{T^{\alpha/2}}. \quad (3)$$

Lemma 3 $\partial_i \leq \sum_{k < i} \Delta_k$.

Proof

Induction on i . Basis follows from the definitions. Step:

$$\begin{aligned} \partial_{i+1} &= |\tilde{V}_i(\xi_0) - V^i(\xi_0)| = |V_i^*(\tilde{V}_{i-1}(\xi_0)) - V_i(V^{i-1}(\xi_0))| \leq \\ &\leq |V_i^*(\xi_i) - V_i(\xi_i)| + |V_i(\xi_i) - V_i(\xi'_i)| = \Delta_i + \partial_i. \end{aligned}$$

Applying the inductive hypothesis we complete the proof.

Thus in view of (3) Lemma 3 gives

$$\forall i = 1, \dots, t \quad \partial_i \leq \frac{2it^{1/2}}{T^{\alpha/2}}. \quad (4)$$

It follows from the definition of the functions f_i that $\forall i \leq t \quad \delta_{x_t}(\xi_i) < \frac{1}{T^\alpha}$. Taking into account inequality (4), we conclude that for $x = x_t$

$$d_x(\xi_i - \xi'_i) \leq \frac{2it^{1/2}}{T^{\alpha/2}}, \quad d_x(\xi_i) < \frac{1}{T^{\alpha/2}}, \quad d_x(\xi'_i) \leq d_x(\xi_i - \xi'_i) + d_x(\xi_i).$$

Hence we have

$$d_x(\xi'_i) \leq \frac{3t^{3/2}}{T^{\alpha/2}}. \quad (5)$$

Now we can change the value of the function f_t only on the word x_t and obtain a new function ϕ such that $\phi^{\{T\}}(\bar{0}) \neq f_t^{\{T\}}(\bar{0})$. Therefore, if $\xi_0 \rightarrow \xi_1'' \rightarrow \dots \rightarrow \xi_t''$ is the computation of $\phi^{\{T\}}(\bar{0})$ on our QC with oracle for ϕ , then we have

$$|\xi'_t - \xi_t''| \geq 1/4. \quad (6)$$

On the other hand, Lemma 2 and inequality (5) give

$$|\xi'_t - \xi_t''| < 2 \sum_{i \leq t} d_x(\xi'_i) \leq \frac{6t^{5/2}}{T^{\alpha/2}} < 1/4$$

for $\alpha \geq 5$ and sufficiently large n , which contradicts to (6). Theorem 1 is proved.

If the time complexity of classical computation exceeds $O(2^{n/7})$ we can establish a lower bound for the time of quantum simulation as $\Omega(T^{1/2})$.

Theorem 2 *For arbitrary function $T(n)$ there is no such QC with oracle for f that for some function $t(n) : t^2/T \rightarrow 0$ ($n \rightarrow \infty$) QC computes $f^{\{T\}}(\bar{0})$ for every f applying only t evaluations of f .*

Proof

Suppose that $8t < \sqrt{T}$. Let f be selected as above. Put $f^k = f^{\{k\}}(\bar{0})$ $k = 0, 1, \dots, T$. Define the matrix $A = (a_{ij})$ with the following elements: $a_{ij} = \delta_{f^j}(\chi_i)$, $i = 0, 1, \dots, t$; $j = 0, 1, \dots, T$. We have for every $i = 0, \dots, t$ $\sum_{j=0}^T a_{ij} \leq 1$, consequently $t \geq \sum_{i=0}^t \sum_{j=0}^T a_{ij} = \sum_{j=0}^T \sum_{i=0}^t a_{ij}$ and there exists such $\tau \in \{0, 1, \dots, T\}$ that $\sum_{i=0}^t a_{i\tau} \leq \frac{t}{T}$.

Changing the value of f only on the word f^τ we obtain a new function g where $g^{\{T\}}(\bar{0}) \neq f^{\{T\}}(\bar{0})$. Let $\chi_0 \rightarrow \chi'_1 \rightarrow \dots \rightarrow \chi'_t$ be computation on QC with oracle for g . Then we have

$$|\chi_t - \chi'_t| \geq 1/4. \quad (7)$$

On the other hand Lemma 2 gives $|\chi_t - \chi'_t| \leq 2 \sum_{i=0}^t \sqrt{a_{i\tau}} \leq 2\sqrt{t \sum_{i=0}^t a_{i\tau}} \leq 2t/T^{1/2} < 1/4$ for sufficiently large n , which contradicts to (7). Theorem 2 is proved.

Note that the lower bound as $\Omega(T^{1/2})$ for the time of quantum simulation follows immediately from the lower bound for the time of quantum search established in the work [BBBV].

5 Acknowledgments

I am grateful to Charles H. Bennett who clarified for me some details of the work [BBBV], to Peter Hoyer for his comments and useful criticism and to Lov K. Grover for his attention to my work.

References

- [BBBV] C.H.Bennett, E.Bernstein, G.Brassard, U.Vazirani, *Strenths and Weaknesses of Quantum Computing*, To appear in SIAM Journal on Computing (lanl e-print quant-ph/9701001)
- [BB] A.Berthiaume, G.Brassard, *Oracle quantum computing*, Journal of modern optics, 41(12):2521-2535, 1994
- [Be] C.H.Bennett, *Logical reversibility of computation*, IBM J. Res.Develop. 17, 525-532

- [DJ] , D.Deutsch, R.Jozsa, *Rapid solution of problems by quantum computation*, Proc. Roy. Soc. Lond. A **439** 553-558
- [FG] E.Farhi, S.Gutmann, *Quantum Mechanical Square Root Speedup in a Structured Search Problem*, lanl e-print, quant-ph/9711035
- [Gr] L.K.Grover, *A fast quantum mechanical algorithm for database search*, Proceedings, STOC 1996, Philadelphia PA USA, pp 212-219
- [Sh] P.W.Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on Quantum Computer*, lanl e-print, quant-ph/9508027 v2 (A preliminary version in Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, Nov. 20-22, 1994, IEEE Computer Society Press, pp 124-134)

Quantum Resonance for Solving NP-complete Problems by Simulations

Michail Zak

Ultracomputing Group, Mail Stop 525-3660
Jet Propulsion Laboratory
California institute of Technology
Pasadena, CA 91109-8099

Abstract. Quantum analog computing is based upon similarity between mathematical formalism of a quantum phenomenon and phenomena to be analyzed. In this paper, the mathematical formalism of quantum resonance combined with tensor product decomposability of unitary evolutions is mapped onto a class of NP-complete combinatorial problems.

1. Introduction

The competition between digital and analog computes, i.e., between computations and simulations, has a long history. During the last fifty years the theory of computation has been based implicitly upon classical physics as idealized in the deterministic Turing machine model. However, despite the many successes of digital computers, the existence of so called hard problems has revealed limitations on their capabilities, since the computational time for solving such problems grows exponentially with the size of the problem.

It was well understood that one possible way to fight the “curse” of the combinatorial explosion is to enrich digital computers with analog devices. In contradistinction to a digital computer, which performs operations on numbers symbolizing an underlying physical process, an analog computer processes information by exploiting physical phenomena directly. It is this problem solving via direct simulation that allows an analog approach to reduce the complexity of the computations significantly. This idea was stressed by Feynman [1] who demonstrated that the problem of exponential complexity in terms of calculated probabilities can be reduced to a problem of polynomial complexity in terms of simulated probabilities. Conceptually, a similar approach can be applied to the whole class of NP-complete problems. But is it possible, in general, to find a new mathematical formulation for any intractable problem in such a way that it becomes tractable? Some experts in computational complexity believe that, in the spirit of the Godel theorem, there always exist computational problems such that every mathematical formulation that captures the essence of the problem is intractable [2]. At this step, we cannot prove or disprove this statement.

There are remarkably few (actually three) papers in which quantum analog computing is discussed. The first one [3] introduces a hypothetical quantum device (a slot machine) for solving a traveling salesman problem. As shown by the author, such a device, although intellectually appealing, requires an exponentially large number of measurements to get the right answer. The second paper [4] discusses the capacity of a hypothetical quantum perception. In the third paper [5], a concept of quantum recurrent networks combining quantum conventional networks with classical feedback loops was introduced and discussed.

In this paper an attempt is made to exploit combinatorial properties of tensor product decomposability of unitary evolution of many-particle quantum systems for simulating solutions to NP-complete problems, while the reinforcement and selection of a desired solution is executed by quantum resonance.

2. Quantum Resonance.

Consider a quantum system characterized by a discrete spectrum of energy eigenstates subject to a small perturbing interaction, and let the perturbation be switched on at zero time. The Hamiltonian of the system can be presented as a sum of the time-independent and oscillating components:

$$H = H_0 + \varepsilon_0 H_1 \int_{\omega} \xi(\omega) \sin \omega t d\omega \quad \varepsilon_0 \ll 1 \quad (1)$$

where H_0 and H_1 are constant Hermitian matrices, ω is the frequency of perturbations, and $\xi(\omega)$ is the spectral density.

The probability of a transition from state k to q in the first approximation is proportional to the product [6] :

$$P_{kq} \propto |\varphi_k^* H_1 \varphi_q|^2 \left[\frac{\sin \frac{1}{2}(a_{qk} - \omega)t}{a_{qk} - \omega} \right]^2 \quad (2)$$

Here φ_j are the eigenstates of H_0 :

$$H_0 \varphi_j = E_j \varphi_j \quad j = 1, 2, \dots, N \quad (3)$$

where E_j are the energy eigenvalues,

$$\hbar a_{kq} = E_k - E_q, \quad k, q = 1, 2, \dots, N \quad (4)$$

and \hbar is the Planck constant.

The resonance, i.e., a time-proportional growth of the transition probability P_{kq} occurs when $\omega = a_{qk}$:

$$P_{kq} = \frac{\pi \varepsilon_0^2}{\hbar^2} |\varphi_k^* H_1 \varphi_q|^2 \xi^2(\omega) t \quad (5)$$

3. Combinatorial Problems

Combinatorial problems are among the hardest in the theory of computations. They include a special class of so called NP-complete problems which are considered to be intractable by most theoretical computer scientists. A typical representative of this class is a famous traveling-salesman problem (TSP) of determining the shortest closed tour that connects a given set of n points in the plane. As for any of NP-complete problem, here the algorithm for solution is very simple: enumerate all the tours, compute their lengths, and select the shortest one. However, the number of tours is proportional to $n!$ and that leads to exponential growth of computational time as a function of the dimensionality n of the problem, and therefore, to computational intractability.

It should be noticed that, in contradistinction to continuous optimization problems where the knowledge about the length of a trajectory is transferred to the neighboring trajectories through the gradient, here the gradient does not exist, and there is no alternative to a simple enumeration of tours.

The class of NP-complete problems has a very interesting property: if any single problem (including its worse case) can be solved in polynomial time, then every NP-complete problem can be solved in polynomial time as well. But despite that, there is no progress so far in removing a curse of combinatorial explosion: it turns out that if one manages to achieve a polynomial time of computation, then the space or energy grow exponentially, i.e., the effect of combinatorial explosion stubbornly reappears. That is why the intractability of NP-complete problems is being observed as a fundamental principle of theory of computations which plays the same role as the second law of thermodynamics in physics.

At the same time, one has to recognize that the theory of computational complexity is an attribute of a digital approach to computations, which means that the monster of NP-completeness is a creature of the Turing machine. As an alternative, one can turn to an analog device which replaces digital computations by physical simulations. Indeed, assume that one found such a physical phenomenon whose mathematical description is equivalent to that of a particular NP-complete problem. Then, incorporating this phenomenon into an appropriate analog device one can simulate the corresponding NP-complete problem. In this connection it is interesting to note that, at first sight, NP-complete problems are fundamentally different from natural phenomena: they look like man-made puzzles and their formal mathematical framework is mapped into decision problems with yes/no solutions. However, one

should recall that physical laws can also be stated in a “man-made” form: The least time (Fermat), the least action (in modifications of Hamilton, Lagrange, or Jacobi), and the least constraints (Gauss).

In this paper we will describe how to map a combinatorial decision problem into the physical phenomenon of quantum resonance on a conceptual level, without going into details of actual implementations.

Let us turn to the property (5) which can be mapped into several computational problems, and, for the purpose of illustration, choose the following one: given n different items to be distributed over n places; the cost of an β^{th} item put in a γ^{th} place is $\lambda_{\beta}^{(\gamma)}$; in general, the costs can be positive or negative, and there are no restrictions to how many different items can be put at the same place. Find yes/no answer to the following question: is there at least one total cost whose absolute value falls into an arbitrarily given interval.

This problem is typical for optimal design. Since the cost of a particular distribution is expressed by the sum

$$E_j = \sum_{\beta=1}^n \lambda_{\beta}^{(\lambda_{\beta})}, \quad j = 1, 2, \dots, N = n^n \quad (6)$$

classically one has to compute all the n^n sums (8) in order to find is there at least one E_q such that

$$a_1 \leq |E_q| \leq a_2, a_2 > a_1 \quad (7)$$

where a_1 and a_2 are arbitrarily prescribed positive numbers.

Since costs $\lambda_{\beta}^{(\gamma)}$ can be positive or negative, the absolute value in Eq. (7) represents a global constraint, and therefore our problem belongs to the class of so called constraint satisfaction problems which are the hardest among other optimization problems. The constraint (7) prevents one from decomposing the solution into smaller-size sub-problems. As shown by Andre Stechert^[7], this problem can be mapped into the partition problem^[8], and therefore, it is NP-complete.

Now we will demonstrate how this problem can be solved by the quantum device described above in one computational step.

First, let us represent the unitary matrix U_0 corresponding to the time-independent Hamiltonian

$$U_0 = e^{iH_0 t} \quad (8)$$

as a tensor product of n diagonal unitary matrices of the size $n \times n$:

$$U_0 = U_1 \otimes U_2 \otimes \dots \otimes U_n \quad (9)$$

where

$$U_\gamma = \begin{pmatrix} e^{i\lambda_1^{(\gamma)}} & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & e^{i\lambda_n^{(\gamma)}} \end{pmatrix} \quad (10)$$

Then the unitary matrix U_0 in (9) will be also diagonal and

$$H_0 = \begin{pmatrix} E_1 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & E_N \end{pmatrix}, \quad N = n^n \quad (11)$$

while E_j is expressed by Eq. (6).

Hence, if one select $\lambda_\beta^{(\gamma)}$ in (8) as the costs of a β^{th} item put in a γ^{th} place, then the eigenstates E_j of the Hamiltonian H_0 will represent costs of all $N = n^n$ possible distributions (8).

Now we have to choose the perturbation of the Hamiltonian, (see Eq. (1)) For that purpose assume that initially the quantum device is in a certain base state k , whose energy E_k does not belong to the interval (7), i.e.,

$$|E_k| < a_1, \text{ or } |E_k| > a_2 \quad (12)$$

and select H_1 and $\xi(\omega)$ as follows:

$$H_1 = P \quad (13)$$

where

$$P = \begin{pmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{pmatrix} \quad (14)$$

and

$$\xi(\omega) = \begin{cases} \xi_0 = Const & \text{if } \frac{|E_k - a_2|}{\hbar} \leq \omega < \frac{|E_k - a_1|}{\hbar} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

Here, for the sake of concreteness, the initial state E_k was selected such that:

$$|E_k - a_1| > |E_k - a_2| \quad (16)$$

Suppose that the given interval a_1, a_2 contains at least one total cost $|E_q|$ from the set (6), i.e., $|E_q|$ satisfies the inequality (7). Then, according to Eqs. (5) and (14), the resonance transition from the initial state E_k to the state E_q (or other states satisfying (7)) will occur with the probability one. Indeed, in the presence of a resonance, the probability for non-resonance transitions are vanishingly small if $\varepsilon_0 \ll 1$ (see Eq. (1)).

However, if the given interval a_1, a_2 does not contain any costs $|E_q|$ from the set (6), then according to Eqs. (5) and (14), there will be no resonance transitions at all, and therefore, with the probability one the quantum device will stay in the initial state.

Thus, in one computational step, the problem is solved in a deterministic way. As follows from Eq. (5), the time required for probability of the resonance transition from the state k to q to become close to one has the order:

$$t^* \sim 0 \left(\frac{\hbar^2}{\varepsilon_0^2 \xi^2 |\varphi_k^* H_1 \varphi_q|^2} \right) \quad (17)$$

4. Conclusion

Thus, it has been demonstrated how a “man-made” problems of exponential computational complexity which is hard to handle by algorithmic methods are solved by exploiting a strongly pronounced physical phenomena: quantum resonance.

The main advantage of the proposed approach is in exponential speedup of solutions to NP-complete combinatorial problems. Two fundamental physical phenomena contribute to it: quantum resonance and tensor-product decomposability of the underlying unitary matrix.

Quantum resonance allows one to represent all the possible solutions to the problem as a set of competing dynamical processes: energy exchanges between pairs of quantum eigenstates. The mathematical formalism of quantum resonance provides a storage for these processes: the transition matrix P_{kq} (see Eq. (2)) where each process is labeled through the corresponding transition probability.

Tensor-product decomposability is a fundamental property of the Schrodinger equation for multi-particle systems. Due to its effect, the number of stored solutions, i.e., the number of transitions P_{kq} is exponentially larger than the number of the

input parameters (see Eq. (6)) and that is what directly contributes into exponential speedup and capacity.

In order to make these two physical phenomena work together, one has to choose the Hamiltonian of the quantum system such that the optimal solution is the winner in the competition with other solutions, i.e., that its transition probability is the largest. This is achieved by selecting the oscillating part of the Hamiltonian in the form of (14).

It should be emphasized that the solution of one NP-complete problem opens up a way to solve every NP-complete problem in polynomial time.

Acknowledgements

The research described in this paper was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology and was sponsored by the National Aeronautics and Space Administration, Office of Space Access and Technology. The author thanks T. Beth, L. Levitin, C. Williams, A. Stechert, and N. Cerf for useful discussions.

References

1. R. Feynman, Int. J. of Theoretical Physics, Vol. 21, No. 6/7 1982.
2. J. Taub and H. Wozniakowski, "Breaking Intractability," Scientific American, August 1992, Vol. 46.
3. V. Cerny, Phys. Rev. A, Vol. 48, No. 1, p. 116, 1993.
4. M. Lewenstein, Quantum Perceptions, J. Modern Optics, 1999, Vol. 41, No. 12 p 2491-8501.
5. M. Zak, C. Williams, Quantum Neural Nets, Int. J. of Theor. Physics. No. 2, 1998.
6. D. R. Bates, Quantum Theory, Acad. Press, N.Y. 1961, p. 271.
7. Andre Stechert, Private Communications, JPL, Jan. 25, 1998
8. M. R. Garey and D. S. Johnson, Computers and Intractability, 1979 W. H. Freeman and Co., NY, p. 90

Computational Complexity and Physical Law^{*}

Daniel S. Abrams¹ and Seth Lloyd²

¹ Department of Physics, MIT 12-128b Cambridge, MA 02139
(abrams@mit.edu)

² d'Arbeloff Laboratory for Information Sciences and Technology
Department of Mechanical Engineering, MIT 3-160 Cambridge, MA 02139
(slloyd@mit.edu)

Abstract. We consider the relationship between computational complexity and the laws of physics, and show that nonlinear generalizations of quantum mechanics allow for polynomial-time solution of NP-complete and #P problems. Whether or not experiments ultimately reveal small nonlinearities in the evolution of quantum states (indeed, we believe that they most likely will not), these results underscore the close relationship between physics and computation.

Not too long ago, it was generally believed that computational complexity was a concept independent of any particular computing model. According to the strong form of Church's thesis, any reasonable computing device could be simulated by any other, using only polynomial resources. The degree to which this point of view was accepted is made evident by the degree of surprise which greeted the counterexample: the discovery that quantum computers could solve in polynomial time problems thought to be classically intractable, such as factoring large numbers.[\[1\]](#)

Yet perhaps the failure of Church's thesis in its strong form should not have been so surprising. After all, the idea is built upon elaborate and non-obvious proofs that one computing model can simulate another. And since a computer is in the end a physical device - whose behavior is governed by the laws of physics - it is only natural that the power of a machine should depend in large part upon the physics which underlies its operation. All the recent work in quantum computing and quantum information has now made this concept quite evident.

In this paper, we explore the possibility of another physical model of computation - one that relies upon nonlinear quantum mechanics. The idea that under some circumstances the superposition principle of quantum mechanics might be violated - that is, that the time evolution of quantum systems might be slightly nonlinear - has been previously suggested in various contexts [\[2\]](#) [\[3\]](#) [\[4\]](#) [\[5\]](#) [\[6\]](#). Such nonlinearity is purely hypothetical: all known experiments confirm the linearity of quantum mechanics to a high degree of accuracy [\[7\]](#) [\[8\]](#) [\[9\]](#) [\[10\]](#). Further, nonlinear

^{*} This work supported in part by grant # N00014-95-1-0975 from the Office of Naval Research, and by ARO and DARPA under grant # DAAH04-96-1-0386 to QUIC, the Quantum Information and Computation initiative, and by a DARPA grant to NMRQC, the Nuclear Magnetic Resonance Quantum Computing initiative.

quantum theories have often been controversial and frequently have had major theoretical difficulties¹ [11] [12] [13]. Nevertheless, the implications of nonlinear quantum mechanics on the theory of computation are profound. In particular, we explain why one can exploit nonlinear time evolution so that the classes of problems NP and #P (including oracle problems) may be solved in polynomial time. Additional algorithms and further details, including an explicit demonstration of how to construct nonlinear quantum logic gates using the Weinberg model, are available in [14].

Loosely defined, the class NP contains all problems for which it is possible to verify a potential solution in polynomial time. These include all problems in the class P (those that can be solved in polynomial time) as well as the NP-complete problems, e.g., the traveling salesman, satisfiability, and sub-graph isomorphism, for which no known polynomial time algorithms exist. One natural way to approach these problems on a quantum computer is to create a superposition of every possible potential solution, and then try to determine if one of those potential solutions is in fact a true solution. In some sense, this technique nicely mimics the theoretical behavior of a non-deterministic Turing machine. In order to both simplify and generalize the result, it is convenient to replace the actual NP problem with an oracle problem, stated as follows: consider an oracle, or “black box”, which calculates a function that maps n bits into a single bit; i.e., it takes an input between 0 and $2^n - 1$ and returns either 0 or 1. We need to determine if there exists an input value x for which $f(x) = 1$. It is easy to see that a polynomial time algorithm to solve this problem can be used to solve all problems in the class NP. (Note, however, that the converse is not necessarily true - the NP complete problems contain structure, whereas the function defined above is completely arbitrary. Thus this oracle problem is in fact a harder problem than those in NP, because it clearly requires exponential time on a classical Turing machine.) For simplicity we will restrict ourselves to the case where there is at most one value x for which $f(x)=1$.

¹ It may be objected that many if not all models of nonlinear quantum mechanics allow for superluminal communication - and thereby backward-in-time signaling, which would then allow in some peculiar sense for the solution of all problems in zero or even negative time. (Although they would still require an exponential number of calls to the oracle). From the authors' perspective, this fact is simply evidence that nonlinear quantum mechanics is most likely wrong. Indeed, we feel that the results described here should probably also be viewed as evidence that quantum mechanics is linear, rather than as a blueprint for a potential device. However, we would like to point out in addition that the “backward-in-time” computer relies upon a combination of relativity and nonlinear quantum mechanics, whereas the algorithms which we describe here and elsewhere do not. That is, they could run on a non-relativistic nonlinear quantum computer. In other words, these algorithms rely upon a fundamentally different aspect of nonlinear quantum mechanics. We do not know if it is possible to create a self-consistent - that is, causal - theory of nonlinear quantum mechanics. However, we do know this: in the end, only experiment can tell what are the correct physical laws.

One might attempt to solve this oracle problem on an ordinary quantum computer using the following technique. First, create a superposition of all the input states:

$$\psi = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i, 0\rangle \quad (1)$$

Next, use the oracle to calculate $f(i)$ for each $|i\rangle$ in parallel:

$$\psi = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i, f(i)\rangle \quad (2)$$

Although the final qubit in some sense “knows” the solution to the problem, a measurement of this qubit will yield $|1\rangle$ with either zero probability if there is no solution or an exponentially small probability if there is a solution. It is therefore necessary to enhance the amplitude of the $|1\rangle$ component of the superposition by an exponentially large factor, in order to distinguish the two cases. One idea is to try to increase the number of states with a $|1\rangle$ rather than increase the amplitude of the particular state $|i\rangle$ for which $f(i) = 1$. Imagine comparing the states $|i\rangle$ in pairs, according to the last bit of $|i\rangle$. Looking at just the last bit of $|i\rangle$ and the final qubit $f(i)$, we see one of the following states:

$$\begin{aligned} (a) \quad & |00\rangle + |11\rangle \\ (b) \quad & |01\rangle + |10\rangle \\ (c) \quad & |00\rangle + |10\rangle \end{aligned} \quad (3)$$

The last case occurs most frequently. What we’d like to do is map these states into new states using the following transformation:

$$\begin{aligned} (a) \quad & |00\rangle + |11\rangle \longrightarrow |01\rangle + |11\rangle \\ (b) \quad & |01\rangle + |10\rangle \longrightarrow |01\rangle + |11\rangle \\ (c) \quad & |00\rangle + |10\rangle \longrightarrow |00\rangle + |10\rangle \end{aligned} \quad (4)$$

This transformation is like an AND gate - it ignores the first qubit and places the second qubit in the state $|1\rangle$ if and only if either of the original components had the state $|1\rangle$ for the second qubit. Performing this transformation on the superposition of all $|i\rangle$ will leave every state unaffected except the state which neighbors the solution $|x\rangle$. This state will then pick up a $|1\rangle$ in place of the $|0\rangle$ which it originally had. If we then compare states in pairs according to the second bit of $|i\rangle$, the number of states with a $|1\rangle$ for the final qubit will double again. Repeated application of this process would then leave the final qubit

unentangled with the first n qubits: it would be either in the pure state $|0\rangle$ if there are no solutions to the problem, or the pure state $|1\rangle$ if there had existed some state $|x\rangle$ for which $f(x)=1$. A measurement of this qubit thereby reveals the answer to the problem.

Of course, this transformation cannot be accomplished using an ordinary quantum computer, because it is nonlinear. That this is the case can be easily seen by the fact that in cases (a) and (c) the initial states are non-orthogonal, but the final states are orthogonal. Hence the desired transformation cannot possibly be linear. One is tempted to try to patch this problem in a variety of ways. One possibility is to imbed this transformation in a larger Hilbert space and hope that a projective subspace might reduce to the desired nonlinear transformation. Unfortunately, this approach cannot succeed. The reason is that different elements of the superposition need to interfere with each other in later stages of the algorithm. If the “linearized” version of the transformation results in extraneous “garbage” qubits, these will prevent the states from interfering with each other in future iterations. Equivalently, one might hope that the non-unitary evolution associated with the measurement process might suffice to accomplish the necessary transformation, but this will of course fail for the same reason. One can also try to hide the extraneous information in the phases of the states. Although this appears promising at first, more careful analysis reveals essentially the same difficulties.²

We see, therefore, that the potential application of a nonlinear quantum logic gate arises naturally from a fairly straightforward approach to the NP oracle problem. From an intuitive perspective, however, it is not exactly clear why it is that nonlinearity is important, beyond the fact that the gate which we desire for our algorithm does not happen to be linear. We can get a better feeling for this from a slightly different perspective.

Consider the shortest-path version of the traveling salesman problem, and a classical algorithm that finds “pretty good” solutions, such as simulated annealing. Implement this algorithm on a quantum computer, and initialize the quantum computer in a state which as before is a superposition of all possible inputs. After the algorithm has finished, the result will be a quantum computer that exists in a superposition of all the various local minima that are found by

² One can also imagine approaching the problem with phases from the very beginning, thereby avoiding the need for the nonlinear gate. After calculating $f(i)$, multiply the phase of the solution state by -1 and then reverse the computation of $f(i)$. The computer would then be in an equal superposition of all $|i\rangle$, with the state $|x\rangle$ for which $f(x)=1$ having opposite phase. Pairs of states containing two $|i\rangle$ of opposite phase are orthogonal to those for which both $|i\rangle$ are of the same phase, so it is possible to reverse the phase of the pair, thereby transferring the minus sign from the solution $|x\rangle$ to its partner state. Repeating the process which created the phase in the first place would then leave two states with negative phase. By iterating through each bit of $|i\rangle$ (as in the previous algorithm), one can continue the process until a substantial fraction of the states have negative phase. This situation can be easily detected. Unfortunately, each iteration takes twice as long as the previous one, so the algorithm described in this footnote requires exponential time.

searching from every possible initial state. A measurement would then reveal any one of these local minima, but most likely not the shortest path. Thus, what one would like to do before the measurement is to compare the various states with each other and shift the amplitude into the states representing shorter paths. Put differently, we would like an algorithm which acts in a space that is restricted to only those quantum states which already have non-zero amplitude. Unfortunately, the linear transformations allowed by ordinary quantum mechanics have no way for a given state (or more precisely, component of a superposition) to “sense” the amplitude of other states. This is the aspect of nonlinear quantum mechanics which allows for the solution of NP-complete problems.

Returning to the algorithm described above, it is clear that if one could obtain the necessary nonlinear transformation, one could find the answer to an NP-complete problem in polynomial (in fact, linear) time, and using only a single evaluation of the oracle. It may be objected that the nonlinear operator described above appears arbitrary and unnatural: indeed, it was selected exactly so as to be able to solve the stated problem. However, the apparently arbitrary operation can be built using ordinary unitary operations and much simpler and more ‘natural’ single qubit nonlinear operators (that is, to the extent that any nonlinear operation in quantum mechanics can be considered ‘natural’). One possible technique for generating the transformation would be to use the following steps: first, act on the two qubits with the unitary operator

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \quad (5)$$

This transforms the states above as follows

$$\begin{aligned} (a) \quad & \frac{1}{\sqrt{2}} [|00\rangle + |11\rangle] \longrightarrow |00\rangle \\ (b) \quad & \frac{1}{\sqrt{2}} [|01\rangle + |10\rangle] \longrightarrow |01\rangle \\ (c) \quad & \frac{1}{\sqrt{2}} [|00\rangle + |10\rangle] \longrightarrow \frac{1}{2} [|00\rangle + |01\rangle - |10\rangle + |11\rangle] \end{aligned} \quad (6)$$

Next, operate on the second qubit with a simple one qubit nonlinear gate \hat{n}_- that maps both $|0\rangle$ and $|1\rangle$ to the state $|0\rangle$. Thus

$$\begin{aligned} (a) \quad & \frac{1}{\sqrt{2}} [|00\rangle + |11\rangle] \longrightarrow |00\rangle \\ (b) \quad & \frac{1}{\sqrt{2}} [|01\rangle + |10\rangle] \longrightarrow |00\rangle \\ (c) \quad & \frac{1}{\sqrt{2}} [|00\rangle + |10\rangle] \longrightarrow |A\rangle \end{aligned} \quad (7)$$

The third final state is unknown because we have not bothered to specify how the non-linear gate acts on the state $|00\rangle + |01\rangle - |10\rangle + |11\rangle$. This omission thereby allows for flexibility in choosing the gate \hat{n}_- . Whatever the state $|A\rangle$ may be, we can perform a unitary operation that will transform the first qubit into the pure state $|0\rangle$ while leaving the state $|00\rangle$ in place. The computer is then in one of the following states

$$\begin{aligned} (a) & |0\rangle|0\rangle \\ (b) & |0\rangle|0\rangle \\ (c) & |0\rangle(x|0\rangle + y|1\rangle) \end{aligned} \tag{8}$$

A second non-linear gate \hat{n}_+ is now required that will map the state $x|0\rangle + y|1\rangle$ to the state $|1\rangle$ (for the particular values of x and y which result from the above steps but not necessarily for arbitrary x and y), while leaving the state $|0\rangle$ unchanged. After this gate is applied, the transformation resulting from the steps described so far is then:

$$\begin{aligned} (a) & \frac{1}{\sqrt{2}} [|00\rangle + |11\rangle] \longrightarrow |00\rangle \\ (b) & \frac{1}{\sqrt{2}} [|01\rangle + |10\rangle] \longrightarrow |00\rangle \\ (c) & \frac{1}{\sqrt{2}} [|00\rangle + |10\rangle] \longrightarrow |01\rangle \end{aligned} \tag{9}$$

The desired two qubit transformation is then easily obtained with a NOT gate on the second qubit and a $\pi/2$ rotation on the first qubit.

Having thus shown how to generate the needed two qubit gate, the question is now reduced to that of generating the simpler single qubit gates \hat{n}_- and \hat{n}_+ . If one considers the state of a qubit as a point on the unit sphere, then all unitary operations correspond to rotations of the sphere; and while such rotations can place two state vectors in any particular position on the sphere, they can never change the angle between two state vectors. A nonlinear transformation corresponds to a stretching of the sphere, which will in general modify this angle. The desired gates \hat{n}_- and \hat{n}_+ are two particular examples of such operations. Excepting perhaps certain pathological cases (e.g., discontinuous transformations), it is evident that virtually any nonlinear operator, when used repeatedly in combination with ordinary unitary transformations (which can be used to place the two state vectors in an arbitrary position on the sphere), can be used to arbitrarily increase or decrease the angle between two states, as needed to generate the gates \hat{n}_- and \hat{n}_+ . We show explicitly in [14] how one can generate these gates using the Weinberg model.

To conclude: we have seen that nonlinear quantum mechanics enables one to solve NP complete problems in polynomial time. In fact, it is not difficult to

generalize this algorithm to solve the $\#P$ oracle problem³, which in turn implies that a nonlinear quantum computer can solve in polynomial time all problems in the entire polynomial hierarchy. Nevertheless, the authors would like to emphasize that these results are probably best viewed as further evidence that the universe is exactly linear - rather than as blueprints for the design of a machine if it were not. (Although it is certainly not obvious, a priori, that quantum mechanics need be strictly linear - and the question can be fairly viewed as an experimental one). Regardless, the connection between physics and computation is now made unavoidable: the underlying laws of physics strongly impact the theoretical complexity of computational problems.

D.S.A. acknowledges support from a NDSEG fellowship, and helpful discussions with J. Jacobson, I. Singer, S. Johnson, I. Park, and T. Wang. Portions of this research were supported by grant # N00014-95-1-0975 from the Office of Naval Research, and by ARO and DARPA under grant # DAAH04-96-1-0386 to QUIC, the Quantum Information and Computation initiative, and by a DARPA grant to NMRQC, the Nuclear Magnetic Resonance Quantum Computing initiative.

References

1. P. Shor, in Proceedings of the 35th Annual Symposium on Foundations of Computer Science, edited by S. Goldwasser (IEEE Computer Society, Los Alamos, CA, 1994), p.124
2. S. Weinberg, Phys. Rev. Lett. 62, 485 (1989)
3. S. Weinberg, Ann. of Phys. 194, pg. 336 (1989)
4. Phys. Rev. A 56, p. 146-56 (1997)
5. B.G. Levy, Physics Today, 12 pp. 20 (1989)
6. O. Bertolami, Physics Letters A, 154, p. 225-9 (1991)
7. P.K. Majumder et. al., Phys. Rev. Lett. 65, 2931 (1990)
8. R.L. Walsworth et. al., Phys. Rev. Lett. 64, 2599 (1990)
9. T.E. Chupp and R.J. Hoare, Phys. Rev. Lett. 64, 2261 (1990)
10. J.J. Bollinger, D.J. Heinzen, W.M. Itano, S.L. Gilbert, D.J. Wineland, Phys. Rev. Lett. 63, 1031 (1989)
11. A. Peres, Phys. Rev. Lett. 63, 1114 (1989)
12. J. Polchinski, Phys. Rev. Lett. 66, pg. 397 (1991)
13. N. Gisin, Phys. Lett. A 113, p. 1 (1990)
14. D.S. Abrams and S. Lloyd, preprint quant-ph/9801041, submitted to Phys. Rev. Lett.

³ To solve the problems in the class $\#P$, one needs to determine the exact number of solutions i for which $f(i)=1$, instead of just whether or not such solutions exist. This can be accomplished if one replaces the flag qubit with a string of $\log_2 n$ qubits and modifies the algorithm slightly - so that it adds the number of solutions in each iteration rather than performing what is effectively a one bit AND. In this case, a measurement of the final result reveals the exact number of solutions.

The Hidden Subgroup Problem and Eigenvalue Estimation on a Quantum Computer

Michele Mosca^{1,2} and Artur Ekert¹

¹ Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, U.K.

² Mathematical Institute, 24-29 St. Giles', Oxford, OX1 3LB, U.K.

Abstract. A quantum computer can efficiently find the order of an element in a group, factors of composite integers, discrete logarithms, stabilisers in Abelian groups, and *hidden* or *unknown* subgroups of Abelian groups. It is already known how to phrase the first four problems as the estimation of eigenvalues of certain unitary operators. Here we show how the solution to the more general Abelian *hidden subgroup problem* can also be described and analysed as such. We then point out how certain instances of these problems can be solved with only one control qubit, or *flying qubits*, instead of entire registers of control qubits.

1 Introduction

Shor's approach to factoring [Sh], (by finding the order of elements in the multiplicative group of integers mod N , referred to as \mathbf{Z}_N^*) is to extract the period in a superposition by applying a Fourier transform. Another approach, based on Kitaev's technique [Ki], is to estimate an eigenvalue of a certain unitary operator. The difference between the two analyses is that the first one considers (or even 'measures' or 'observes') the *target* or *output* register in the standard computational basis, while the analysis we detail here considers the target register in a basis containing eigenvectors of unitary operators related to the function f . The actual network of quantum gates, as highlighted in [CEMM], is the same for both algorithms; it is helpful to understand both approaches. In some cases, which we discuss in Sect. 5, this approach suggests implementations which do not require a register of control qubits. A more general formulation of the order-finding problem as well as the discrete logarithm problem, and the Abelian stabiliser problem, is the *hidden subgroup problem* (or the *unknown* subgroup problem [Ho]). In the case that G is presented as the product of a finite number of cyclic groups (so G is finitely generated and Abelian), all of these problems are solved by the familiar sequence of a Fourier transform, a function application, and an inverse Fourier transform. In this paper we describe how this more general problem can also be viewed and analysed as an estimation of eigenvalues of unitary operators.

2 The Hidden Subgroup Problem

Let f be a function from a finitely generated group G to a finite set X such that f is constant on the cosets of a subgroup K (of finite index, since X is finite), and distinct on each coset. The hidden subgroup problem is to find K (that is, a generating set for K), given a way of computing f . When K is normal in G , we could in fact decompose f as $h \circ g$, where g is a homomorphism from G to some finite group H , and h is some 1-to-1 mapping from H to the set X . In this case, K corresponds to the kernel of g and H is isomorphic to G/K . We will occasionally refer to this decomposition, which we illustrate in Fig. [1](#). Define the *input size*,

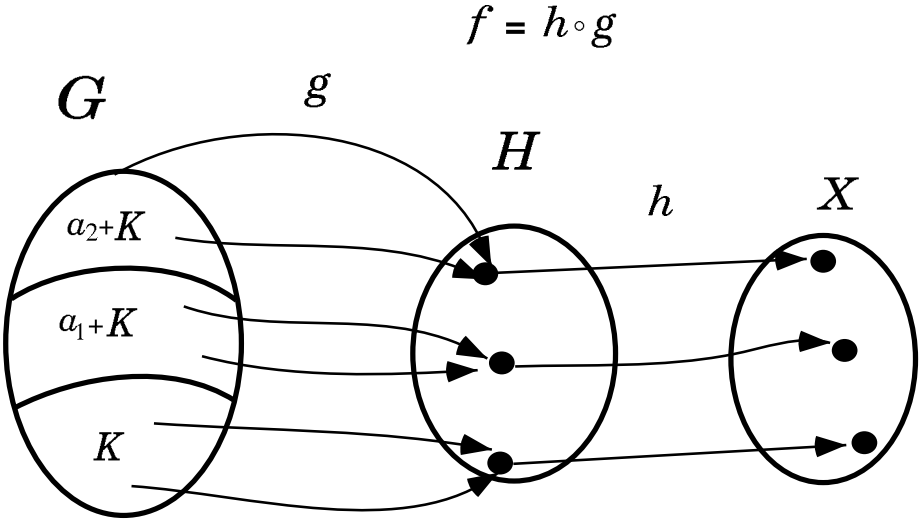


Fig. 1. The function f can be viewed as the composition of a homomorphism g to a group H , and some 1-to-1 mapping h to the set X . Our hidden subgroup K will be the kernel of g , and H is isomorphic to G/K .

n , to be of order $\log_2[G : K]$. We will count the number of operations, or the *running time*, in terms of n . An algorithm is considered *efficient* if its running time is polynomial in the input size. By *elementary quantum operations*, we are referring to a finite set of quantum logic gates which allow us to approximate any unitary operation. See [\[BBCDMSSW\]](#) for a discussion and further references. Our running times will always refer to expected running times, unless explicitly stated otherwise. By *expected running time* we are referring to the expected number of operations for *any* input (and not just an average of the expected running times over all inputs).

We should be clear about what it means to have a finitely generated group G , and to be able to compute the function f . This is difficult without losing

some generality or being dry and technical, or both. The algorithms we describe only apply for groups G which are represented as finite tuples of integers corresponding to the direct product of cyclic groups (consequently, G is finitely generated and Abelian). Conversely, for any finitely generated Abelian G , there is a temptation to point out that G is isomorphic to such a direct product of cyclic groups, and *assume* that we can easily access this product structure. This is not always the case, even in cases of practical interest. For example, \mathbf{Z}_N^* , the multiplicative group of integers modulo N for some large integer N , which is Abelian of order $\phi(N)$ (the Euler ϕ -function) and thus isomorphic to a product of cyclic groups of prime power order. We will not necessarily know $\phi(N)$ or have a factorisation of it along with a set of generators for \mathbf{Z}_N^* . However, in light of the quantum algorithms described in this paper, we could efficiently find such an isomorphism, thereby increasing the number of finitely generated Abelian groups which can be efficiently expressed in a manner which allows us to employ these algorithms. We will however leave further discussion of these details to another note [EM]. When we talk about computing f , we assume that we have some unitary operation U_f which takes us from state $|\mathbf{x}\rangle|0\rangle$ to $|\mathbf{x}\rangle|f(\mathbf{x})\rangle$. It could, for example, take $|\mathbf{x}\rangle|\mathbf{y}\rangle$ to $|\mathbf{x}\rangle|\mathbf{y} + f(\mathbf{x})\rangle$, where $+$ denotes an appropriate group operation, such as addition modulo N when the second register is used to represent the integers modulo N .

Various cases of the hidden subgroup problem are described in [Si], [Sh], [Ki], [BL], [Gr], [Jo], [CEMM], and [Hø]. We note that [BL] also covers the case that f is not necessarily distinct on each coset (that is, h is not 1-to-1), and this is discussed in the appendix. Finding the order r of an element in a group H of unknown size, or the period r of a function f , is a special case where $G = \mathbf{Z}$ and $K = r\mathbf{Z}$. For any generator \mathbf{e}_j of a finitely generated G , we can use the algorithm in Sect. 4.2 to find an integer k such that $f(k\mathbf{e}_j) = f(0)$, so that $k\mathbf{e}_j \in K$. We find this k with $O(n)$ applications of f and $O(n^2)$ other elementary quantum operations. We can then assume that \mathbf{e}_j is of order k (that is, factor $\langle k\mathbf{e}_j \rangle$ out of G), and in general assume that G is a finite group.

We give a few examples.

Deutsch's Problem: Consider a function f mapping $\mathbf{Z}_2 = \{0, 1\}$ to $\{0, 1\}$. Then $f(x) = f(y)$ if and only if $x - y \in K$, where K is either $\{0\}$ or $\mathbf{Z}_2 = \{0, 1\}$. If K is $\{0\}$, then f is 1-to-1 (or *balanced*), and if K is \mathbf{Z}_2 then f is constant. [De, CEMM]

Simon's Problem: Consider a function f from \mathbf{Z}_2^l to some set X with the property that $f(x) = f(y)$ if and only if $x - y \in \{0, \mathbf{s}\}$ for some string \mathbf{s} of length l . Here $K = \{0, \mathbf{s}\}$ is the hidden subgroup of \mathbf{Z}_2^l . Simon [Si] presents an efficient algorithm for solving this problem, and the solution to the hidden subgroup problem in the Abelian case is a generalisation.

Discrete Logarithms: Let G be the group $\mathbf{Z}_r \times \mathbf{Z}_r$ where \mathbf{Z}_r is the additive group of integers modulo r . Let the set X be the subgroup generated by some element a of a group H , with $a^r = 1$. For example, $H = \mathbf{F}_q^*$, the multiplicative group of the field of order q , where $r = q - 1$. Let $a, b \in G$, and suppose

$b = a^m$. Define f to map (x, y) to $a^x b^y$. Here the hidden subgroup of G is $K = \{(k, -km) | k = 0, 1, \dots, r-1\} = \langle (1, -m) \rangle$, the subgroup generated by $(1, -m)$. Finding this hidden subgroup will give us the logarithm of b to the base a . The security of the U.S. Digital Signature Algorithm is based on the computational difficulty of this problem in \mathbf{F}_q^* (see [MOV] for details and references). Here the input size is $n = \lceil \log_2 r \rceil$. Shor's algorithm [Sh] was the first to solve this problem efficiently. In this case, f is also a homomorphism which can make implementations more simple as described in Sect. 5.

Self-Shift-Equivalent Polynomials: Given a polynomial P in l variables X_1, X_2, \dots, X_l over \mathbf{F}_q , the function f which maps $(a_1, a_2, \dots, a_l) \in \mathbf{F}_q^l$ to $P(X_1 - a_1, X_2 - a_2, \dots, X_l - a_l)$ is constant on cosets of a subgroup K of \mathbf{F}_q^l . This subgroup K is the set of self-shift-equivalences of the polynomial P . Grigoriev [Gr] shows how to compute this subgroup. He also shows, in the case that q has characteristic 2, how to decide if two polynomials P_1 and P_2 are shift-equivalent, and to generate the set of elements (a_1, a_2, \dots, a_l) such that $P_1(X_1 - a_1, X_2 - a_2, \dots, X_l - a_l) = P_2(X_1, X_2, \dots, X_l)$. The input size n is at most $l \log_2 q$.

Abelian Stabiliser Problem: Let G be any group acting on a finite set X . That is, each element of G acts as a map from X to X , in such a way that for any two elements $a, b \in G$, $a(b(x)) = (ab)(x)$ for all $x \in X$. For a particular element x of X , the set of elements which fix x (that is, the elements $a \in G$ such that $a(x) = x$), form a subgroup. This subgroup is called the stabiliser of x in G , denoted $St_G(x)$. Let f_x denote the function from G to X which maps $g \in G$ to $g(x)$. The hidden subgroup corresponding to f_x is $K = St_G(x)$. The finitely generated Abelian case of this problem was solved by Kitaev [Ki], and includes finding orders and discrete logarithms as special cases.

3 Phase Estimation and the Quantum Fourier Transform

In this section, we review the relationship between phase estimation and the quantum Fourier transform which was highlighted in [CEMM].

The quantum Fourier transform for the cyclic group of order N , F_N , maps

$$|a\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{2\pi i ax/N} |x\rangle.$$

So F_N^{-1} maps

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{2\pi i ax/N} |x\rangle \rightarrow |a\rangle.$$

More generally, for any ϕ , $0 \leq \phi < 1$, F_N^{-1} maps

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} e^{2\pi i \phi x} |x\rangle \rightarrow \sum_{x=0}^{N-1} \alpha_{\phi, x} |x\rangle \quad (1)$$

where the amplitudes $\alpha_{\phi,x}$ are concentrated near values of x such that x/N are good estimates of ϕ . The closest estimate of ϕ will have amplitude at least $4/\pi^2$. The probability that x/N will be within k/N of ϕ is at least $1 - 1/(2k - 1)$. See [CEMM] for details in the case that N is a power of 2; the same proof works for any N . Thus to estimate ϕ such that, with probability at least $1 - \epsilon$, the error is less than $1/M$, we should use a control register containing values from 0 to $N - 1$ and apply F_N^{-1} for any $N \geq M(1/\epsilon + 1)/2$. For example, if we desire an error of at most $1/2^n$ with probability at least $1 - 1/2^m$ we could use $N = 2^{n+m}$. In practice, it will be best to use the N that corresponds to the group that is easiest to represent and work with in the particular physical realisation of the quantum computer at hand. We expect that this N will be a power of two.

For convenience, we will omit normalising factors in the remainder of this paper. It will also be convenient to have a compact notation for the state on the right hand side of (II) which we consider to be a good estimator for $|\phi\rangle$. So let us refer to this state as $|\tilde{\phi}\rangle_N$ or just $|\tilde{\phi}\rangle$ if the value of N is understood. Lastly, we will use $\exp(x)$ to denote e^x .

4 The Algorithm

To restrict attention from finitely generated groups G to finite groups we need to know how to solve the cyclic case (just one generator), that is, to find the period of a function from \mathbf{Z} to the set X . We will first describe how to find the order of an element a in a group H , or equivalently, the period of the function $f : t \rightarrow a^t$, as Shor [Sh] did for the group $H = \mathbf{Z}_N^*$, the multiplicative group of integers modulo N . We will then show how to generalise it to find the period of any function $f : \mathbf{Z} \rightarrow X$. If f were a homomorphism (so h is an isomorphism of H , when f is decomposed as $f = h \circ g$), we would just be finding the order of $f(1)$ in H . The difference is that we are showing how to deal with a non-trivial h which hides the homomorphism structure. The details will also help explain how to find hidden subgroups of finite Abelian groups.

4.1 Finding Orders in Groups

We have an element a from a group H and we wish to find the smallest positive integer r such that $a^r = 1$. The group H is not necessarily Abelian; all that matters is that the subgroup generated by a is Abelian, and this is always true. The idea is to create an operator U_a which corresponds to multiplication by a (so it maps $|y\rangle$ to $|ay\rangle$). Since $a^r = 1$, then $U_a^r = I$, the identity operator. Hence the eigenvalues of U_a are r th roots of unity, $\exp(2\pi i k/r)$, $k = 0, 1, \dots, r - 1$. By estimating a random eigenvalue of U_a , with accuracy $1/2r^2$, we can determine the fraction k/r . The denominator (with the fraction in lowest terms) will be a factor of r . We thus seek to estimate an eigenvalue of U_a ; note that $U_a^r = U_{a^r}$.

For any integer x define U_{a^x} to be the operator that maps $|y\rangle$ to $|a^x y\rangle$. Define U_{a^x} to be the operator which maps $|x\rangle|y\rangle$ to $|x\rangle U_{a^x}|y\rangle = |x\rangle|a^x y\rangle$. Note that U_{a^x} acts on two registers and x is a variable which takes on the value

in the first register, while U_{a^x} acts on one register and x is fixed. Consider the eigenvectors

$$|\Psi_k\rangle = \sum_{t=0}^{r-1} \exp(-2\pi i k t / r) |a^t\rangle, k = 0, 1, \dots, r-1, \quad (2)$$

of U_{a^x} and respective eigenvalues $\exp(2\pi i k x / r)$. If we start with the superposition

$$\sum_{x=0}^{2^l-1} |x\rangle |\Psi_k\rangle$$

and then apply U_{a^x} we get

$$\sum_{x=0}^{2^l-1} \exp(2\pi i k x / r) |x\rangle |\Psi_k\rangle.$$

As discussed in the previous section, applying $F_{2^l}^{-1}$ to the first register gives $|\widetilde{k/r}\rangle |\Psi_k\rangle$ and thus a good estimate of k/r .

We will not typically have $|\Psi_k\rangle$ but we do know that $|1\rangle = \sum_{k=0}^r |\Psi_k\rangle$. Therefore we can start with

$$|0\rangle |1\rangle = |0\rangle \sum_{k=0}^r |\Psi_k\rangle = \sum_{k=0}^r |0\rangle |\Psi_k\rangle \quad (3)$$

and then apply F_{2^l} to the first register to produce

$$\sum_{k=0}^{r-1} \left(\sum_{x=0}^{2^l-1} |x\rangle \right) |\Psi_k\rangle. \quad (4)$$

We then apply U_{a^x} to get

$$\sum_{k=0}^{r-1} \left(\sum_{x=0}^{2^l-1} \exp(2\pi i k x / r) |x\rangle \right) |\Psi_k\rangle \quad (5)$$

followed by $F_{2^l}^{-1}$ on the control register to yield

$$\sum_{k=0}^{r-1} |\widetilde{k/r}\rangle |\Psi_k\rangle. \quad (6)$$

Observing the first register will give an estimate of k/r for an integer k chosen uniformly at random from the set $\{0, 1, \dots, r-1\}$. As shown in [Sh], we choose $l > 2 \log_2 r$, and use the continued fractions algorithm to find the fraction k/r . Of course, we do not know r , so we must either use an l we know will be larger than $2 \log_2 r$, such as $2 \log_2 N$ in the case that H is \mathbf{Z}_N^* . (Alternatively, we could guess

a lower bound for r , and if the algorithm fails, subsequently double the guess and repeat.) We then repeat $O(1)$ times to find r . This algorithm thus uses $O(1)$ exponentiations, or $O(n)$ group multiplications, and $O(n^2)$ elementary quantum operations to do the Fourier transforms.

We can factor the integer N by finding orders of elements in \mathbf{Z}_N^* . This uses only $O(n^3)$ or $\exp(c \log n)$ elementary quantum operations, for $c = 3 + o(1)$ (or $c = 2 + o(1)$ if we use fast Fourier transform techniques). Other deterministic factoring methods will factor N in $O(\sqrt{N})$ or $\exp(cn)$ steps, where $c = 1/2 + o(1)$. The best known rigorous probabilistic classical algorithm (using index calculus methods) [LP] uses $\exp(c(n \log n)^{1/2})$ elementary classical operations, $c = 1 + o(1)$. There is also an algorithm with a heuristic expected running time of $\exp(c(n^{1/3}(\log n)^{2/3}))$ elementary classical operations (see [MOV] for an overview and references) for $c = 1.902 + o(1)$. Thus, in terms of elementary operations, a quantum computer provides a drastic improvement over known classical methods to factor integers.

4.2 Finding the Period of a Function

The above algorithm, as pointed out in [BL], can be applied to a more general setting. Replace the mapping from t to a^t with any function f from the integers to some finite set X . Define $U_{f(x)}$ to be an operator that maps $f(y)$ to $f(y+x)$. This is a generalisation of U_{a^x} except it does not matter how it is defined on values not in the range of f , as long as it is unitary. Define $U_{f(x)}$ to be an operator which maps $|x\rangle |f(y)\rangle$ to $|x\rangle U_{f(x)} |f(y)\rangle = |x\rangle |f(y+x)\rangle$.

The following are eigenvectors of $U_{f(x)}$:

$$|\Psi_k\rangle = \sum_{t=0}^{r-1} \exp(-2\pi i k t / r) |f(t)\rangle, \quad k = 0, 1, \dots, r-1, \quad (7)$$

with respective eigenvalues $\exp(2\pi i k x / r)$. As in (3), we can start with

$$|0\rangle |f(0)\rangle = \sum_{k=0}^{r-1} |0\rangle |\Psi_k\rangle$$

except with our new, more general, definition of $|\Psi_k\rangle$. We apply F_{2^n} to the first register to produce (4), and then apply $U_{f(x)}$ to produce (5), followed by $F_{2^n}^{-1}$ to get (6). Observing the first register will give an estimate of k/r for an integer k chosen uniformly at random, and the same analysis as in the previous section applies to find r .

One important issue is how to compute $U_{f(x)}$ only knowing how to compute f . Note that from (4) to (5) (using the modified definition of $|\Psi_k\rangle$) we simply go from

$$\sum_{x=0}^{2^n-1} |x\rangle |f(0)\rangle = \sum_{x=0}^{2^n-1} \left(\sum_{k=0}^{r-1} |x\rangle |\Psi_k\rangle \right) \quad (8)$$

to

$$\sum_{x=0}^{2^n-1} |x\rangle |f(x)\rangle = \sum_{x=0}^{2^n-1} |x\rangle \left(\sum_{k=0}^{r-1} \exp(2\pi i x k / r) |\Psi_k\rangle \right) \quad (9)$$

which could be accomplished by applying U_f , which we do have, to the starting state

$$\sum_{x=0}^{2^n-1} |x\rangle |0\rangle.$$

Thus even if we do not know how to explicitly compute the operators $U_{f(x)}$, any operator U_f which computes the function f will give us the state (9). This state permits us to estimate an eigenvalue of $U_{f(x)}$ which lets us find the period of the function f with just $O(1)$ applications of the operator U_f and $O(n^2)$ other elementary operations. The equality in (9) is the key to the equivalence between the two approaches to these quantum algorithms. On the left hand side is the original approach ([Si], [Sh], [BL]) which considers the target register in the standard computational basis. We can analyse the Fourier transform of the preimages of these basis states, which is less easy when the Fourier transforms do not exactly correspond to the group G . On the right hand side of (9) we consider the target register in a basis containing the eigenvectors of the unitary operators which we apply to it (as done in [Ki] and [CEMM], for example), and this gives us (5), from which it is easy to see and analyse the effect of the inverse Fourier transform even when it does not perfectly match the size of G .

4.3 Finding Hidden Subgroups

As discussed in Sect. 2, any finite Abelian group G is the product of cyclic groups. In light of the order-finding algorithm, which also permits us to factor, we can assume that the group G is represented as a product of cyclic groups of prime power order. Further, for any product of two groups G_p and G_q whose orders are coprime, any subgroup K of $G_p \times G_q$ must be equal to $K_p \times K_q$ from some subgroups K_p and K_q of G_p and G_q respectively. We can therefore consider our function f separately on G_p and G_q and determine K_p and K_q separately. Thus we can further restrict ourselves to groups G of prime power order. This not only simplifies any analysis, it could reduce the size of quantum control registers necessary in any implementation of these algorithms.

Let us thus assume that $G = \mathbf{Z}_{p^{m_1}} \times \mathbf{Z}_{p^{m_2}} \times \cdots \times \mathbf{Z}_{p^{m_l}}$ for some prime p and positive integers $m_1 \leq m_2 \leq \cdots \leq m_l = m$. Let $K = \{\mathbf{k} = (k_1, k_2, \dots, k_l) | f(\mathbf{x}) = f(\mathbf{x} + \mathbf{k}) \text{ for all } \mathbf{x} \in G\}$. The 'promise' is that f is constant on cosets of K , and distinct on each coset. In practice, this will usually be a consequence of the nature of f , as in the case of discrete logarithms where $f(x_1, x_2) = a^{x_1} b^{x_2}$, or whenever f is constructed as $h \circ g$ for some homomorphism g from G to some finite group H , and a 1-to-1 mapping h from H to the set X .

Let U_f be an operator which maps $|\mathbf{x}\rangle |0\rangle$ to $|\mathbf{x}\rangle |f(\mathbf{x})\rangle$. Define $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, and so on. Let us also consider an operator related to U_f , $U_{f(\mathbf{x}\mathbf{e}_j)}$, which maps $|x\rangle |f(\mathbf{y})\rangle$ to $|x\rangle U_{f(\mathbf{x}\mathbf{e}_j)} |f(\mathbf{y})\rangle = |x\rangle |f(\mathbf{y} + \mathbf{x}\mathbf{e}_j)\rangle$. In the

case of Simon's Problem, $U_{f(x(0,1,0))}$ maps $|1\rangle |f(y_1, y_2, y_3)\rangle$ to $|1\rangle U_{f(0,1,0)} |f(y_1, y_2, y_3)\rangle = |1\rangle |f(y_1, y_2 + 1, y_3)\rangle$ and does nothing to $|0\rangle |f(y_1, y_2, y_3)\rangle$.

For each $\mathbf{t} = (t_1, t_2, \dots, t_l)$, $0 \leq t_j < p^{m_j}$, satisfying

$$\sum_{j=1}^l p^{m-m_j} h_j t_j = 0 \pmod{p^m} \quad \text{for all } \mathbf{h} \in K \quad (10)$$

define

$$|\Psi_{\mathbf{t}}\rangle = \sum_{\mathbf{a} \in G/K} \exp\left(\frac{-2\pi i}{p^m} \sum_{j=1}^l p^{m-m_j} t_j a_j\right) |f(\mathbf{a})\rangle. \quad (11)$$

We are summing over a set of representatives of the cosets of K modulo G , and by condition (10) on \mathbf{t} , this sum is well-defined. Let T denote the set of \mathbf{t} satisfying (10), which corresponds to the group of characters of G/K . The $|\Psi_{\mathbf{t}}\rangle$ are eigenvectors of each $U_{f(x\mathbf{e}_j)}$, with respective eigenvalues $\exp(2\pi i x t_j / p^{m_j})$. By determining these eigenvalues, for $j = 1, 2, \dots, l$, we will determine \mathbf{t} . If we had $|\Psi_{\mathbf{t}}\rangle$ in an auxiliary register, we could estimate t_j / p^{m_j} using $U_{f(x\mathbf{e}_j)}$ by the technique of the previous section. If we use $F_{p^{m_j}}^{-1}$ we would determine t_j exactly, or we could use the simpler $F_{2^k}^{-1}$, for some $k > \log_2(p^{m_j})$, and obtain t_j with high probability. For simplicity, we will use $F_{p^{m_j}}^{-1}$. In practice we could use $F_{2^k}^{-1}$ for a large enough k so that the probability of error is sufficiently small.

By estimating t_j / p^{m_j} for $j = 1, 2, \dots, l$, we determine \mathbf{t} . The algorithm starts by preparing l control registers in the state $|0\rangle$ and one *target* or *auxiliary* register in the state $|\Psi_{\mathbf{t}}\rangle$, applies the appropriate Fourier transforms to produce

$$\left(\sum_{x_1=0}^{p^{m_1}-1} |x_1\rangle\right) \cdots \left(\sum_{x_l=0}^{p^{m_l}-1} |x_l\rangle\right) |\Psi_{\mathbf{t}}\rangle \quad (12)$$

followed by $U_{f(x\mathbf{e}_j)}$ for $j = 1, 2, \dots, n$, using the j th register as the control and $|\Psi_{\mathbf{t}}\rangle$ as the target, to produce

$$\left(\sum_{x_1=0}^{p^{m_1}-1} \exp(2\pi i \frac{x_1 t_1}{p^{m_1}}) |x_1\rangle\right) \cdots \left(\sum_{x_l=0}^{p^{m_l}-1} \exp(2\pi i \frac{x_l t_l}{p^{m_l}}) |x_l\rangle\right) |\Psi_{\mathbf{t}}\rangle. \quad (13)$$

Then apply $F_{p^{m_j}}^{-1}$ to the j th control register for each j to yield

$$|t_1\rangle |t_2\rangle \cdots |t_l\rangle |\Psi_{\mathbf{t}}\rangle \quad (14)$$

from which we can extract \mathbf{t} . As in the previous section, we do not know how to construct $|\Psi_{\mathbf{t}}\rangle$, but we do know that

$$|f(\mathbf{0})\rangle = \sum_{\mathbf{t} \in T} |\Psi_{\mathbf{t}}\rangle.$$

So we start with

$$|\mathbf{0}\rangle |\mathbf{0}\rangle \cdots |\mathbf{0}\rangle |f(\mathbf{0})\rangle = \sum_{\mathbf{t} \in T} |\mathbf{0}\rangle |\mathbf{0}\rangle \cdots |\mathbf{0}\rangle |\Psi_{\mathbf{t}}\rangle$$

apply Fourier transforms to get

$$\sum_{\mathbf{t} \in T} \left(\sum_{x_1=0}^{p^{m_1}-1} |x_1\rangle \right) \cdots \left(\sum_{x_l=0}^{p^{m_l}-1} |x_l\rangle \right) |\Psi_{\mathbf{t}}\rangle \quad (15)$$

then apply $U_{f(\mathbf{x}_{\mathbf{e}_j})}$ using the j th register as a control register, for $j = 1, 2, \dots, n$, and the last register as the target register to produce

$$\sum_{\mathbf{t} \in T} \left(\sum_{x_1=0}^{p^{m_1}-1} \exp(2\pi i \frac{x_1 t_1}{p^{m_1}}) |x_1\rangle \right) \cdots \left(\sum_{x_l=0}^{p^{m_l}-1} \exp(2\pi i \frac{x_l t_l}{p^{m_l}}) |x_l\rangle \right) |\Psi_{\mathbf{t}}\rangle. \quad (16)$$

We finally apply $F_{p^{m_j}}^{-1}$ to the j th control register for $j = 1, 2, \dots, l$, to produce

$$\sum_{\mathbf{t} \in T} |\mathbf{t}\rangle |\Psi_{\mathbf{t}}\rangle. \quad (17)$$

Observing the first register lets us sample the \mathbf{t} 's uniformly at random, and thus with $O(n)$ repetitions we will, by (10), have enough independent linear relations for us to determine a generating set for K . For example, in the case of Simon's problem, the $|\mathbf{t}\rangle$ all satisfy $\mathbf{t} \cdot \mathbf{s} = \sum_{j=1}^l t_j s_j \bmod 2 = 0 \bmod 2$, where $K = \{0, \mathbf{s}\}$. We could also guarantee that each new non-zero element of T will increase the span by a technique discussed in the appendix.

This analysis of eigenvectors and eigenvalues is based on the work in [Ki]. The problem is that, unlike in [Ki], we do not always have the operator $U_{f(\mathbf{x}_{\mathbf{e}_j})}$. However, note that, like in Sect. 4.2, going from (15) to (16) maps

$$\left(\sum_{0 \leq x_j \leq p^{m_j}} |\mathbf{x}\rangle \right) |f(0)\rangle$$

to

$$\begin{aligned} & \sum_{0 \leq x_j \leq p^{m_j}} |\mathbf{x}\rangle |f(\mathbf{x})\rangle \\ &= \sum_{\mathbf{t} \in T} \left(\sum_{x_1=0}^{p^{m_1}-1} \exp(2\pi i \frac{x_1 t_1}{p^{m_1}}) |x_1\rangle \right) \cdots \left(\sum_{x_l=0}^{p^{m_l}-1} \exp(2\pi i \frac{x_l t_l}{p^{m_l}}) |x_l\rangle \right) |\Psi_{\mathbf{t}}\rangle. \end{aligned}$$

We can create state (16) by applying U_f , which we do have, to the starting state

$$\sum_{0 \leq x_i < p^{m_i}} |\mathbf{x}\rangle |0\rangle$$

and proceeding with the remainder of the algorithm. As in Sect. 4.2, we are considering the target register in the basis containing the eigenvectors $|\Psi_k\rangle$ instead of the computational basis.

5 Reducing the Size of Control Registers

5.1 Discrete Logarithms

In practice, it might be advantageous to reduce the number of qubits required to solve a problem, or the length of time each qubit must be isolated from the environment. For example, suppose we wish to find m such that $a^m = b$, where the order of a divides r . The operators U_{a^x} and U_{b^x} , which correspond to multiplication by a^x and b^x respectively, share the eigenvectors $|\Psi_k\rangle$ (see [2]) and have corresponding eigenvalues $\exp(2\pi i k x / r)$ and $\exp(2\pi i k m x / r)$. We can assume we know r by applying the order-finding algorithm if necessary. By using U_{a^x} with one control register we can approximate k/r , and by using U_{b^x} with another control register we can approximate $(km \bmod r)/r$ and then extract m modulo $r/\gcd(r, k)$. Note that since we know r , we only need $\log r$ bits of precision when estimating k/r and $(km \bmod r)/r$, instead of $2 \log_2 r$ when using continued fractions. Note further that, knowing r , it may be possible to actually place $|\Psi_k\rangle$ into the target register (by direct construction or otherwise) for some known k , and thus only require *one* control register with over $\log_2 r$ qubits to estimate $(km \bmod r)/r$. One way of doing this is to keep the target register after we have applied the order-finding algorithm and observed an estimate of k/r in the control register. At this point, the target register is almost entirely in the state $|\Psi_k\rangle$, and we could now just estimate the eigenvalue of U_{b^x} on this eigenstate, which we know will be $(km \bmod r)/r$.

5.2 One Control Bit

Consider the case that we have an efficient computational means of mapping $|f(\mathbf{y})\rangle$ to $|f(\mathbf{y} + \mathbf{x})\rangle$ for any \mathbf{x} . If we consider f to be of the form $h \circ g$ for a homomorphism g , we are requiring that h is the identity or some other function with enough structure that we can efficiently map $h(g(\mathbf{y}))$ to $h(g(\mathbf{y} + \mathbf{x})) = h(g(\mathbf{y}) + g(\mathbf{x}))$. In this case we can efficiently solve the hidden subgroup problem with only one control bit or a sequence of *flying qubits* [THLMK]. We illustrate this method for the problem of finding the order of an element a in a group H .

Figure 2 shows the relationship between $F_{2^n}^{-1}$ and the controlled multiplications by powers of a in the order-finding algorithm. As already pointed out in [GN], the measurements could be performed before the controlled rotations. The quantum controlled rotations could then be replaced with semi-classically controlled rotations of the subsequent qubits. This brings us to Fig. 3, where we observe further that all the operations on the first qubit could be performed before we even prepare the second qubit. All the operations could be done sequentially, starting from the first qubit, the results of measuring the previous qubits determining how to prepare the next qubit before measurement. This means we could in fact do all the quantum controlled multiplications with a single control qubit provided we can execute the semi-classical controls which allow us to reset a qubit to $|0\rangle + |1\rangle$ and perform a rotation dependent upon the previous measurements (the rotations could in fact be implemented at any

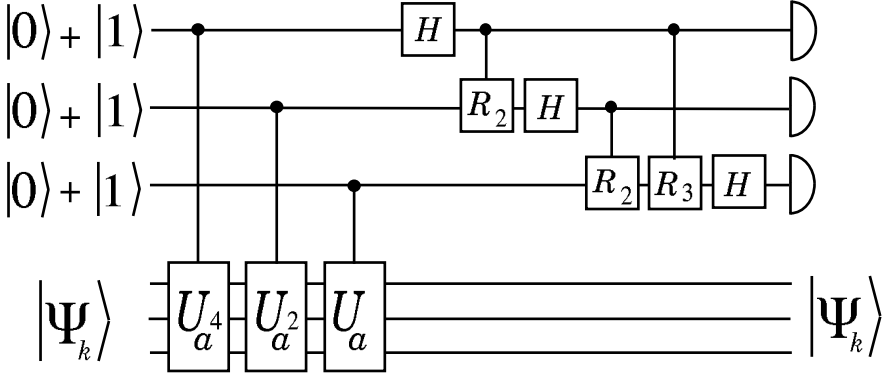


Fig. 2. We start with $(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)|\Psi_k\rangle = \sum_{x=0}^7 |x\rangle |\Psi_k\rangle$. The controlled multiplications create the state $\sum_{x=0}^7 \exp(2\pi i k x/r) |x\rangle |\Psi_k\rangle$. The remaining gates create the state $|\widetilde{k/r}\rangle$ (apart from reversing the order of the qubits) which we then observe. The H -gates correspond to *Hadamard* transforms, and the R_j -gates correspond to a controlled phase shift of $\exp(2\pi i/2^j)$ on state $|1\rangle$.

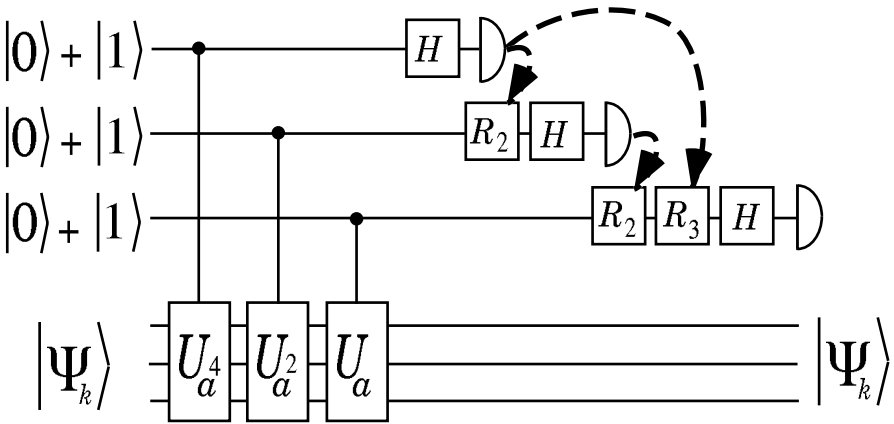


Fig. 3. Here we employ a semi-classical version of F_{23}^{-1} . We could measure each qubit before it is used as a control, perform the controlled rotations semi-classically, and the probability of observing each possible output state $|x_1\rangle |x_2\rangle |x_3\rangle$ is the same as in Fig. 2.

time after resetting the qubit and before applying the final Hadamard transform and measuring it; they could also be omitted provided we repeat each step a few extra times and do some additional classical post-processing as done in [Ki]). Alternatively, the control qubits could be a sequence of flying qubits which are measured (or prepared) in a way dependent upon the outcomes of the previous measurements of control qubits.

For the more general hidden subgroup problem in Abelian groups we would have a sequence of applications of $U_{f(x\mathbf{e}_j)}$ controlled by one qubit, which is measured, then reset to a superposition of $|0\rangle$ and $|1\rangle$ plus some rotation that is dependent upon the previous measurements. In summary,

Remark 1. The hidden subgroup K of a finitely generated Abelian group G generated by $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$, corresponding to a function f from G to a finite set X , can be found with probability close to 1 by semi-classical methods with only *one* control bit (or a sequence of flying qubits) and polynomial in n applications of the operators $|x\rangle |f(\mathbf{y})\rangle \rightarrow |x\rangle |f(\mathbf{y} + x\mathbf{e}_j)\rangle$ for $j = 1, 2, \dots, k$, where n is the index of K in G .

Acknowledgments

Many thanks to Peter Høyer for helping prepare this paper, to Mark Ettinger and Richard Hughes for helpful discussions and hospitality in Los Alamos, to BRICS (Basic Research in Computer Science, Centre of the Danish National Research Foundation), and to Wolfson College.

This work was supported in part by CESG, the European TMR Research Network ERP-4061PL95-1412, Hewlett-Packard, The Royal Society London, and the U.S. National Science Foundation under Grant No. PHY94-07194. Part of this work was done at the 1997 Elsag-Bailey – I.S.I. Foundation workshop on quantum computation, at NIS-8 division of the Los Alamos National Laboratory, and at the BRICS 1998 workshop on Algorithms in Quantum Information Processing.

References

- [BBCDMSSSW] Barenco, A., Bennett, C.H., Cleve, R., DiVincenzo, D.P., Margolus, N., Shor, P., Sleater, T., Smolin, J., Weinfurter, H.: *Phys. Rev. A* **52**, (1995) 3457.
- [BL] Boneh, D., and Lipton, R.J.: Quantum cryptanalysis of hidden linear functions (Extended abstract). *Lecture Notes on Computer Science* **963** (1995) 424–437
- [CEMM] Cleve, R., Ekert, E., Macchiavello, C., and Mosca, M.: Quantum Algorithms Revisited, *Proc. Roy. Soc. Lond. A*, **454**, (1998) 339–354.
- [De] Deutsch, D. : Quantum Theory, the Church-Turing principle and the universal quantum computer. *Proc. Roy. Soc. Lond. A*, **400**, (1985) 97–117.
- [EM] Ekert, A., Mosca, M.: (note in preparation, 1998).
- [GN] Griffiths, R.B. and Niu, C.-S.: Semi-classical Fourier Transform for Quantum Computation, *Phys. Rev. Lett.* **76** (1996) 3228–3231.

- [Gr] Grigoriev, D. Y.: Testing the shift-equivalence of polynomials by deterministic, probabilistic and quantum machines. Theoretical Computer Science, **180** (1997) 217-228.
- [Hø] Høyer, P. : Conjugated Operators in Quantum Algorithms. preprint, (1997).
- [Jo] Jozsa, R.: Quantum Algorithms and the Fourier Transform, Proc. Roy. Soc. Lond. A, **454**, (1998) 323-337.
- [Ki] Kitaev, A. Y. : Quantum measurements and the Abelian stabiliser problem. e-print quant-ph/9511026 (1995)
- [LP] Lenstra, H. W. Jr., and Pomerance, C.: A Rigorous Time Bound For Factoring Integers, Journal of the AMS, Volume 5, Number 2, (1992) 483-516.
- [MOV] Menezes, A., van Oorschot, P., Vanstone, S. : Handbook of Applied Cryptography, C.R.C. Press, 1997.
- [Sh] Shor, P. : Algorithms for quantum computation: Discrete logarithms and factoring. Proc. 35th Ann. Symp. on Foundations of Comp. Sci. (1994) 124-134
- [Si] Simon, D.: On the Power of Quantum Computation. Proc. 35th Ann. Symp. on Foundations of Comp. Sci. (1994) 116-123
- [THLMK] Turchette, Q.A., Hood C.J., Lange W., Mabuchi H., and Kimble H.J.: Measurement of conditional phase shifts for quantum logic. Phys. Rev. Lett. , **76**, 3108 (1996).

Appendix: When f Is Many-to-1 on G/K

The question of what happens when f is many-to-1 on cosets of K was first addressed in [BL]. This is a slight weakening of the promise that f is distinct on each coset. Suppose f can have up to m cosets going to the same output, for some known m . That is, $f = h \circ g$ where g is a homomorphism from G to a some group H with kernel K , and h is a mapping from H to X that is at most m -to-1. If m divides the order of K , we clearly have a problem. For example, suppose K is the cyclic group of order $2M$, and $m = 2$, but by changing one value of f it would have period M . It can easily be shown that $\Omega(\sqrt{M})$ (that is, at least $c\sqrt{M}$ for some positive constant c) applications of f are necessary to distinguish such a modified f from the original one with probability greater than $3/4$, and thus no polynomial time algorithm, quantum or classical, could distinguish the two cases. Thus one requirement for there to exist an efficient solution in the worst case is that m is less than the smallest prime factor of $|K|$, the number of elements in K .

The problem when f is not 1-to-1 is the following. Running the same quantum algorithm will produce the state

$$\sum_{k=0}^{r-1} |\widetilde{k/r}\rangle |\Psi'_k\rangle$$

where

$$|\Psi'_k\rangle = \sum_{t=0}^{r-1} \exp(-2\pi i k t / r) |f(t)\rangle.$$

This is the same definition as in [7] except now the $|f(t)\rangle$ are not necessarily distinct. This means the sizes of each of the $|\Psi'_k\rangle$ are not necessarily the same

since both destructive and constructive interference can occur. Also, the $|\Psi'_k\rangle$ are no longer orthogonal, and thus some constructive interference could occur on the poor estimates of k/r . Recall that even the close estimates of k/r will not yield useful results when $k = 0$. Any other k will at least reveal a small factor of r . So we need to guarantee that the probability of observing a close enough estimate of k/r for some $k \neq 0$ is significant.

By making our estimates precise enough, say by using over $2\log_2 r + \epsilon/m^2$ control qubits, the estimates of k/r will have error less than $1/2r^2$ (so that continued fractions will work) with probability at least $1 - \epsilon/m^2$. Thus assuming f is 1-to-1, the probability of observing a *bad* output other than 0 would be at most ϵ/m^2 , and the probability of observing 0 would be at most $1/r + \epsilon/m^2$. However, since f is at most m -to-1, these probabilities could amplify by at most a factor of m^2 to ϵ and $m^2/r + \epsilon$ respectively. Observing a 0 means we either got a bad output, or the period of f is 1. Getting 0 as a bad output is not very harmful, however getting another bad output is more complicated, since it will give us a false factor of r . It will be useful to make ϵ small, so that it is unlikely our answer is tainted by false factors of r . Once we have one factor r_1 of r , we can replace $f(x)$ with $f(r_1x)$ (as done in [BL]), which has period r/r_1 and find a factor of r/r_1 . Once we have a big enough factor r' of r , we might start observing 0's, which tells us that the remaining factor of the original r , namely r/r' , is less than m^2 . Thus we can explicitly test $f(r'), f(2r'), f(3r'), \dots$, until we find the period, which will occur after at most m^2 applications. We thus have an algorithm with running time, in terms of elementary quantum operations and applications of f , polynomial in $\log(r)$ and quadratic in m .

The trick of reducing the order of the function can be applied to reduce the size of the group and hidden subgroup in the finite Abelian hidden subgroup problem. When $G = \mathbf{Z}_p$, we can efficiently test if $K = G$ or $K = \{1\}$. The above analysis tells us how to deal with the case that $G = \mathbf{Z}_{p^n}$ for $n > 1$. A similar technique will reduce $G = \mathbf{Z}_{p^{t_1}} \times \dots \times \mathbf{Z}_{p^{t_k}}$ to a quotient group \overline{G} and we can again proceed inductively until the size of \overline{G} is less than m^2 . We can then exhaustively test \overline{G} for the hidden subgroup \overline{K} in another $O(m^2)$ steps.

We emphasize that this is a worst-case analysis. If there were a noticeable difference in the behaviour of a 1-to-1 and an m -to-1 function f , $m > 1$, we could decide if a given function h is 1-to-1 or many-to-one (by composing h with a function f whose period or hidden Abelian subgroup we know, and test for this difference in behaviour). Distinguishing 1-to-1 functions from many-to-1 functions seems like a *very* difficult task in general, and would solve the graph automorphism problem, for example.

A Diakoptic Approach to Quantum Computation

Giuseppe Castagnoli¹ and Dalida Monti²


¹ Information Technology Dept., Elsag Bailey, 16154 Genova, Italy

² Università di Genova and Elsag Bailey, 16154 Genova, Italy

Abstract. In the diakoptic approach, mechanisms are divided into simpler parts interconnected in some standard way (say by a “mechanical Connection”). We *explore* the possibility of applying this approach to quantum mechanisms; the specialties of the quantum domain seem to yield a richer result. First parts are made independent of each other by assuming that Connections are removed. The overall state would thus become a superposition of tensor products of the eigenstates of the independent parts. Connections are restored by projecting off all the tensor products which violate them. This would be performed by particle statistics, under a special interpretation thereof. The NP-complete problem of testing the satisfiability of a Boolean network is approached in this way. The diakoptic approach appears to be potentially able of taming the quantum whole without clipping its richness.

PACS: 89.70.+c, 89.80.+h.

1 Definition of Quantum Mechanical Connection

In (classical) applied mechanics, the diakoptic (dissectionistic) approach is exemplified by the notion of mechanical Connection. Connections divide the whole into simpler parts and reconstruct it — they introduce a “divide and conquer” strategy. In fig. 1(a), a crank-shaft is the Connection which imposes an invertible function between the positions of parts r and s (here discretized as 0 and 1, then the function is the Boolean NOT) 

Things can be more difficult in quantum mechanics, since the Hamiltonian of a Connection may not commute with the parts Hamiltonians. This difficulty is avoided by implementing each Connection through a form of constructive and destructive interference, assumedly related to particle statistics. By applying reverse engineering, the Connection is first introduced as a mathematical feature that would be nice-to-have in quantum mechanisms. Then we ask ourselves whether that feature can be physical.

Let us consider the mechanism of fig. 1 from a quantum perspective. The Connection should establish a constraint between two otherwise independent quantum parts r and s , with eigenstates respectively $|0\rangle_r, |1\rangle_r$ and $|0\rangle_s, |1\rangle_s$ (fig. 1b). Their overall state should have the form

¹ The term “diakoptic” is borrowed from Gabriel Kron, who has developed a powerful formalization of this engineering methodology^[1].

$$|\varphi\rangle = \alpha |0\rangle_r |1\rangle_s + \beta |1\rangle_r |0\rangle_s, \text{ with } |\alpha|^2 + |\beta|^2 = 1;$$

the eigenvalues of each tensor product satisfy the Boolean NOT (the constraint) and $|\varphi\rangle$ is free to “move” in the two-dimensional Hilbert space $H_s = \text{span}\{|0\rangle_r |1\rangle_s, |1\rangle_r |0\rangle_s\}$ – which gives the one degree of freedom required from a Connection.

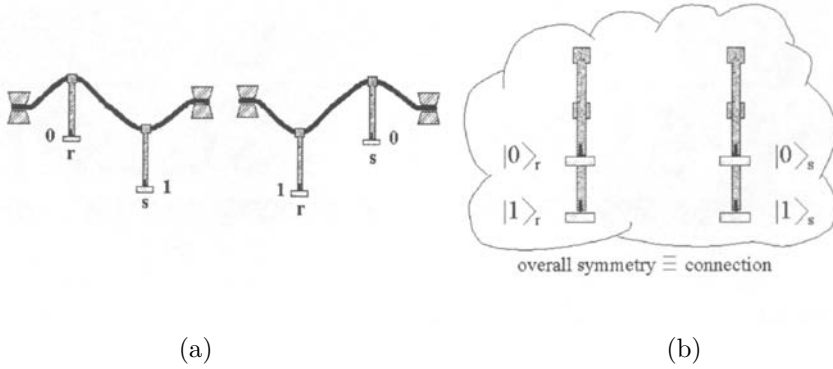


FIGURE 1.

Let us assume the Connection to be temporarily removed. The generic states of the two independent parts are: $|\Psi\rangle_r = \alpha_r |0\rangle_r + \beta_r |1\rangle_r$, $|\Psi\rangle_s = \alpha_s |0\rangle_s + \beta_s |1\rangle_s$. The whole unentangled state in the qubits Hilbert space H_w is $|\Psi\rangle = \alpha_0 |0\rangle_r |0\rangle_s + \alpha_1 |0\rangle_r |1\rangle_s + \alpha_2 |1\rangle_r |0\rangle_s + \alpha_3 |1\rangle_r |1\rangle_s$, with $\alpha_o = \alpha_r \alpha_s$, etc. The Connection is restored by projecting $|\Psi\rangle$ on the “symmetric” subspace H_s . Let us define the projector (or “symmetry”) A_{rs} by:

$$\begin{aligned} A_{rs} |0\rangle_r |1\rangle_s &= |0\rangle_r |1\rangle_s, \quad A_{rs} |1\rangle_r |0\rangle_s = |1\rangle_r |0\rangle_s, \\ A_{rs} |0\rangle_r |0\rangle_s &= A_{rs} |1\rangle_r |1\rangle_s = 0. \end{aligned}$$

The A_{rs} projection of $|\Psi\rangle$ is the normalized vector of H_s closest to it. This is obtained (in a peculiar way whose motivation will be clarified) by submitting a *free normalized vector* $|\varphi\rangle$ of H_w (whose amplitudes on the basis vectors of H_w are free and independent variables up to normalization) to the mathematically simultaneous conditions: (i) $A_{rs} |\varphi\rangle = |\varphi\rangle$, and (ii) the distance between the vector before projection $|\Psi\rangle$ and that after projection $|\varphi\rangle$ should be minimum; in equivalent terms $\|\langle\Psi|\varphi\rangle\|$ should be maximum. This yields the usual result $|\varphi\rangle = k (\alpha_1 |0\rangle_r |1\rangle_s + \alpha_2 |1\rangle_r |0\rangle_s)$, an allowed Connection state (k is the renormalization factor). The Connection will perform by operating on the *parts* under continuous A_{rs} projection of the *whole* on H_s .

2 A Diakoptic Interpretation of Particle Statistics

To give an introductory example, let us show a sort of Connection simply related to particle statistics. Let 1 and 2 be two free, identical and non-interacting spin $1/2$ particles. At a given time, their overall spatial wave function is obtained by symmetrizing/antisymmetrizing (under particles permutation P_{12}) the product of the two independent wave functions:

$$\Psi(x_1, x_2) \cong (1 \pm P_{12}) \Psi_A(x_1) \Psi_B(x_2) = e^{ik_A x_1} e^{ik_B x_2} \pm e^{ik_A x_2} e^{ik_B x_1};$$

x_1 and x_2 are the particles spatial coordinates; the $+$ ($-$) sign goes with the spin singlet (triplet) state (normalization is disregarded). As readily seen: $\|\Psi(x_1, x_2)\|^2 = \cos^2 kx$ for the singlet state, $\|\Psi(x_1, x_2)\|^2 = \sin^2 kx$ for the triplet state, where $x = x_1 - x_2$, $k = k_A - k_B$. Thus close (separated) particles are more likely to be found in a singlet (triplet) state. There is a sort of Connection inducing a correlation between the mutual distance of the two particles and the character of their spin state. Noticeably, this kind of Connection would fall apart if the two particles were not identical. We should note that symmetrization (antisymmetrization) can be seen as the result of *continuous projection* (at any time t) of the state of the two *independent* particles on a symmetrical (antisymmetrical) subspace.

Let us discuss the notion that particle statistics symmetries can be due to projection. As another example, we consider a pair of identical bosons labeled 1 and 2; $S_{12} = \frac{1}{2}(1 + P_{12})$ is the usual symmetrization projector; 0/1 stand for, say, horizontal/vertical polarization. The symmetry $S_{12}|\Psi\rangle = |\Psi\rangle$ is satisfied in

$$H_t = \text{span} \left\{ |0\rangle_1 |0\rangle_2, |1\rangle_1 |1\rangle_2, \frac{1}{\sqrt{2}}(|0\rangle_1 |1\rangle_2 + |1\rangle_1 |0\rangle_2) \right\}.$$

There is a common didactic way of introducing this kind of symmetry. First, statistics is disregarded and the particles are assumed to be independent of each other (like in the above case of the two free particles). Let their *unentangled* state at time t be $|\Psi(t)\rangle' = \alpha_0 |0\rangle_1 |0\rangle_2 + \alpha_1 |0\rangle_1 |1\rangle_2 + \alpha_2 |1\rangle_1 |0\rangle_2 + \alpha_3 |1\rangle_1 |1\rangle_2$. Second, statistics is recovered by symmetrizing $|\Psi(t)\rangle'$, namely by projecting it on H_t . We generalize the example of the free particles by taking this didactic procedure seriously: particle statistics is interpreted as the result of *projection on a predetermined Hilbert subspace* – the one satisfying the symmetry – of a system state which could otherwise be out of symmetry (as will be clarified in the following Sections).

We shall discuss this interpretation preliminarily. It amounts to considering the equation

$$\forall t : S_{12} |\Psi(t)\rangle = |\Psi(t)\rangle, \quad (2.1)$$

as a *constraint* applied to $|\Psi(t)\rangle$. When a particle statistics symmetry is an initial condition conserved as a constant of motion, this constraint is redundant. However, the notion of Connection will be related to particle statistics by means of a

counterfactual reasoning based on eq. (2.1). The idea is that $|\Psi(t)\rangle$, symmetrical at time t_1 , *could* be pushed out of symmetry at an immediately subsequent time $t_2 > t_1$; but in this case eq. (2.1) would project it back on H_t . Particle statistics would operate like a special watch-dog effect internal to the endosystem, or like destructive and constructive interference, by killing the amplitudes of those basis vectors (in a suitable reference) of $|\Psi(t_2)\rangle$ which violate the symmetry, and reinforcing the other amplitudes through re-normalization. This can be seen as a continuous form of partial state vector reduction on a symmetrical subspace.

To see why $|\Psi(t)\rangle$ could be “pushed out of symmetry”, we consider the system defined in Section 1 and A_{rs} projection. In a first step, A_{rs} projection is disregarded while parts r and s are assumed to be independent of each other. An operation on part r could well push the overall state $|\Psi(t)\rangle$ out of symmetry, but in a second step this is prevented by the continuous projection of $|\Psi(t)\rangle$ on H_s : $\forall t : A_{rs} |\Psi(t)\rangle = |\Psi(t)\rangle$. A_{rs} will be related to A_{12} in Section 5.

We should note that this projection (or, if one prefers, state vector reduction on a predetermined subspace) will in general alter the entanglement between the parts r and s , thus the coherence elements of $\rho_r(t)$ (part r density matrix). However, it does not alter the diagonal elements of $\rho_r(t)$; the diagonal is determined by the operation performed on part r , namely it is a constraint to be satisfied by projection. Furthermore, the distance between the vector before projection and the vector after projection should be minimum. Interestingly, this is like partial state vector reduction — from a Hilbert space H_w to a subspace $H_s \subset H_w$. The outcome of reduction can be obtained by submitting a free vector of H_s to the condition that its distance from the vector before reduction is minimum^[12]. As a matter of fact, it is like particle statistics induced a continuous form of (partial) state vector reduction on a predetermined subspace.

3 Behaviour of the Quantum Mechanical Connection

Given the Connection r,s defined in Section 1, let us consider an operation performed on just one qubit, say r . Let this be the continuous rotation $Q_r(\varphi) = \cos \varphi |0\rangle_r \langle 0|_r - \sin \varphi |0\rangle_r \langle 1|_r + \sin \varphi |1\rangle_r \langle 0|_r + \cos \varphi |1\rangle_r \langle 1|_r$, with $\varphi = \omega t$ and t going from 0 to $\frac{\varphi F}{\omega}$. We shall examine the effect of applying $Q_r(\varphi)$ to qubit r ,

$$\rho_r(t) = Q_r(\omega t) \rho_r(0) Q_r^\dagger(\omega t), \quad (3.1)$$

under continuous A_{rs} projection of the overall state.

Let the Connection initial state be the “symmetrical” state (whose tensor products satisfy symmetry A_{rs}):

$$|\Psi(0)\rangle = \cos \vartheta |0\rangle_r |1\rangle_s + \sin \vartheta |1\rangle_r |0\rangle_s. \quad (3.2)$$

Successive states are obtained by submitting a free normalized vector $|\Psi(t)\rangle$ of the Hilbert space H_w (Section 1) to the *mathematically simultaneous* conditions:

for all t (or φ):

- i) $A_{rs} |\Psi(t)\rangle = |\Psi(t)\rangle$;
- ii) $\rho_r(t) = Tr_s(|\Psi(t)\rangle \langle \Psi(t)|) = \cos^2(\vartheta + \varphi) |0\rangle_r \langle 0|_r + \sin^2(\vartheta + \varphi) |1\rangle_r \langle 1|_r$; Tr_s means partial trace over s . Under condition (i), $|\Psi(t)\rangle$ has the form $\alpha(t) |0\rangle_r |1\rangle_s + \beta(t) |1\rangle_r |0\rangle_s$, thus $\rho_r(t)$ is always a diagonal matrix: its coherent elements are killed by A_{rs} projection (or reduction);
- iii) the distance between the vectors before and after projection is minimum. Since projection is continuous, $\| |\Psi(t)\rangle - |\Psi(t + \Delta t)\rangle \|$ must be maximized orderly for $t = 0, t = \Delta t, t = 2\Delta t, \dots, t = N\Delta t$, where $\Delta t = \frac{\varphi}{N\omega}$; then the limit for $N \rightarrow \infty$ must be taken (however, maximization ordering turns out to be irrelevant).

Conditions (i) and (ii) yield $|\Psi(t)\rangle = \cos(\vartheta + \varphi) |0\rangle_r |1\rangle_s + e^{i\delta} \sin(\vartheta + \varphi) |1\rangle_r |0\rangle_s$, with δ unconstrained, as can be checked; condition (iii), given the initial state (3.2), sets $\delta = 0$, yielding to the unitary evolution ($\varphi = \omega t$):

$$|\Psi(t)\rangle = \cos(\vartheta + \varphi) |0\rangle_r |1\rangle_s + \sin(\vartheta + \varphi) |1\rangle_r |0\rangle_s. \quad (3.3)$$

This makes a “good” Connection. Qubit r rotation is transmitted to s :

$$Tr_r(|\Psi(t)\rangle \langle \Psi(t)|) = \rho_s(t) = \sin^2(\vartheta + \varphi) |0\rangle_s \langle 0|_s + \cos^2(\vartheta + \varphi) |1\rangle_s \langle 1|_s. \quad (3.4)$$

Eigenvalues 0 and 1 are interchanged since one qubit is the NOT of the other. Noticeably, by simultaneously rotating the other extremity s of the Connection by the same amount, the same result (3.3) is obtained. This means adding eq. (3.4) as a condition, but this is redundant with respect to (i) and (ii), it was derived from (i) and (ii). Whereas, two different rotations of the two Connection extremities give an impossible mathematical system; this is a rigid Connection.

It should be noted that a rotation φ of qubit (part) r under A_{rs} projection, is equivalent to applying the unitary operator $Q(\varphi)$ to the overall state $|\Psi(t)\rangle$:

$$Q(\varphi) \equiv \begin{pmatrix} \cos \varphi & \sin \varphi & 0 & 0 \\ -\sin \varphi & \cos \varphi & 0 & 0 \\ 0 & 0 & \cos \varphi & -\sin \varphi \\ 0 & 0 & \sin \varphi & \cos \varphi \end{pmatrix},$$

$$\text{with } |0\rangle_r |1\rangle_s \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, |1\rangle_r |0\rangle_s \equiv \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

It can be seen that $Q(\varphi)$ operates on the overall state in an irreducible way, bringing it from $|\Psi(0)\rangle$ (3.2) to $|\Psi(t)\rangle$ (3.3) without ever violating A_{rs} . We have thus ascertained a peculiar fact. Our operation on a part, *blind* to its effect on the whole, performed together with continuous A_{rs} projection, generates a *unitary transformation* which is, so to speak, *wise* to the whole state, to how it should be transformed without violating A_{rs} . Of course A_{rs} ends up commuting with the resulting overall unitary propagator (shaped by it).

4 Quantum Computation Networks

Let us consider the reversible Boolean network of fig. 2(a), fully deployed in space – time is orthogonal to the network lay-out. This is different from sequential computation, where the Boolean network appears in the computation space-time diagram.

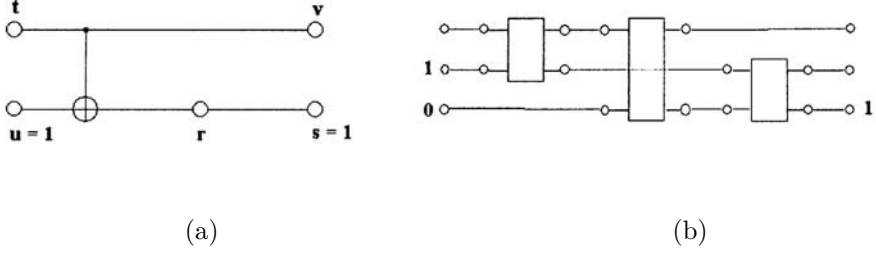


FIGURE 2.

Nodes t , u and v , r make the input and the output of a controlled NOT; r and s belong to a Connection. The c-NOT is made up of the four *coexisting* qubits t , u , v , r ; it has four eigenstates which map the gate Boolean relation and constitute the basis of $H_g =$

$$\text{span}\{|0\rangle_t |0\rangle_u |0\rangle_v |0\rangle_r, |0\rangle_t |1\rangle_u |0\rangle_v |1\rangle_r, |1\rangle_t |0\rangle_u |1\rangle_v |1\rangle_r, |1\rangle_t |1\rangle_u |1\rangle_v |0\rangle_r\}.$$

Model Hamiltonians of such gates are given in [2]; they are different from time-sequential gates where the input and output are successive states of the same register.^[3,4,5,6]

The satisfiability problem is stated by constraining *part* of the input and *part* of the output (just one output qubit is sufficient), and asking whether this network admits a solution. Let $u = 1$ and $s = 1$ be such constraints. $u = 1$ ($s = 1$) propagates a *conditional* logical implication from left to right (right to left). Logical implication is conditioned by the values of the unconstrained part of the input (output). To have a solution, the two propagations must be matched, i.e. they must generate a univocal set of values on all the nodes of the network. Finding whether the network admits at least one match (one solution) is an NP-complete problem. Possible collisions (mismatch) between the two propagations will be both overcome and reconciled by the Connection.

Let us assume that the network has just one solution (which is the case here: $t = 1$, $u = 1$, $r = 0$, $v = 1$, $s = 1$). The procedure to find it is as follows (this will hold for a generic network, thus we can think of many gates and Connections – see fig. 2b, where each wire should incorporate a NOT function and stands for a Connection). The output constraint is removed while an arbitrary value, here $t = 0$, is assigned to the unconstrained part of the input. The logical propagation of this input toward the output yields $t = 0$, $u = 1$, $r = 1$, $s = 0$ ($v = t$ will be

understood). This computation is performed off line in polynomial time (in the network size). It serves to specify the initial state in which the network must be prepared: $|\Psi(0)\rangle = |0\rangle_t |1\rangle_u |1\rangle_r |0\rangle_s$. This state satisfies the gate/s and the Connection/s, but qubit s is in $|0\rangle_s \langle 0|_s$ rather than $|1\rangle_s \langle 1|_s$ (the output constraint). It will be continuously rotated from $|0\rangle_s \langle 0|_s$ to $|1\rangle_s \langle 1|_s$ under A_{rs} projection, while keeping $\rho_u = |1\rangle_u \langle 1|_u$ fixed. This transformation operates on the network Hilbert space H_n ; here $H_n = H_g \otimes H_s$ where $H_s = \text{span}\{|0\rangle_s, |1\rangle_s\}$. Note that all states of H_n natively satisfy the gate/s, not necessarily the Connection/s.

At any time t , the state of the network is obtained by submitting a free normalized state $|\Psi(t)\rangle$ of H_n to the conditions:

for all t :

- i) $A_{rs} |\Psi(t)\rangle = |\Psi(t)\rangle$;
- ii) $Tr_{t,u,r}(|\Psi(t)\rangle \langle \Psi(t)|) = \rho_s(t) = \cos^2 \varphi |0\rangle_s \langle 0|_s + \sin^2 \varphi |1\rangle_s \langle 1|_s$, with $\varphi = \omega t$ and t going from 0 to $\frac{\pi}{2\omega}$;
- iii) $Tr_{t,r,s}(|\Psi(t)\rangle \langle \Psi(t)|) = \rho_u(0) = |1\rangle_u \langle 1|_u$; in a generic network there might be more conditions of this kind;
- iv) the distance between the vectors before and after projection is minimum as specified in Section 3.

This yields:

$$|\Psi(t)\rangle = \cos \varphi |0\rangle_t |1\rangle_u |1\rangle_r |0\rangle_s + e^{i\delta} \sin \varphi |1\rangle_t |1\rangle_u |0\rangle_r |1\rangle_s, \quad (4.1)$$

as is readily checked. For $\varphi = \frac{\pi}{2}$, one obtains $|\Psi(\frac{\pi}{2\omega})\rangle = |1\rangle_t |1\rangle_u |0\rangle_r |1\rangle_s$, namely the solution.

Transformation (4.1) brings the state of the network from satisfying only the input to satisfying both the input and the output constraints. It is obtained by “blindly” operating on *divided parts* of the network, but under A_{rs} projection/s (the *conquering* factor).

If δ does not change with time, it can be seen that propagation (4.1) is unitary. Let us further discuss this point; more generally, we consider a network admitting one or more solutions. The generic vector of H_n satisfying (i) through (iii) can be written sorting out solutions and non-solutions:

$$|\Psi(t)\rangle = \left(\cos \varphi \sum_{i=0}^{L-1} \alpha_i |i\rangle_Q |0\rangle_s + \sin \varphi \sum_{i=L}^{M-1} \alpha_i |i\rangle_Q |1\rangle_s \right) |1\rangle_u \dots,$$

with $\sum_{i=0}^{L-1} \|\alpha_i\|^2 = \sum_{i=L}^{M-1} \|\alpha_i\|^2 = 1$. $|1\rangle_u \dots$ denotes the tensor product of the constrained input qubits; s is the output qubit, whose density operator should be brought to $|1\rangle_s \langle 1|_s$; $|i\rangle_Q$ denotes a tensor product of all other network qubits ($\langle i|_Q |j\rangle_Q = \delta_{i,j}$). The network tensor products corresponding to the first summation do not satisfy the output constraint. Those of the second summation satisfy it.

Let $|0\rangle_Q |0\rangle_s |1\rangle_u \dots$ be the network preparation. At $\varphi \cong 0_+$, “immediately after” starting qubit s rotation and for $0 \leq i \leq L-1$, condition (iv) sets

$\alpha_i = \delta_{0,i}$ whereas, for $L \leq i \leq M-1$, the α_i are left unconstrained but for the above normalization condition. If one takes a random choice of them, this choice must stay frozen throughout $\rho_s(t)$ rotation, because of (iv) as is readily seen. Measurement at $\varphi = \frac{\pi}{2}$ yields a solution $|k\rangle_Q |1\rangle_s |1\rangle_u \dots$ ($L \leq k \leq M-1$). Assuming that state vector reduction can be placed everywhere from preparation to measurement, if one prefers it could be located “immediately before” $\varphi \cong 0_+$. In this case qubit s rotation would rise from the beginning just one, randomly chosen, solution (namely, $\alpha_i = \delta_{k,i}$ for $L \leq \alpha_i \leq M-1$). Anyhow, for $\varphi \geq 0_+$ the network state can be written

$$|\Psi(t)\rangle = \cos \varphi |0\rangle_s |1\rangle_P + e^{i\delta} \sin \varphi |1\rangle_s |0\rangle_P,$$

where $|0\rangle_P, |1\rangle_P$ are two *constant* (in time) orthonormal vectors of H_g . For a constant δ (randomly chosen once for all), $|\Psi(t)\rangle$ is a unitary evolution in the subspace $H_s \otimes H_P$, with $H_P = \text{span}\{|0\rangle_P, |1\rangle_P\}$.

As a result of the foregoing process, A_{rs} symmetries (or projectors) become constants of motion which commute with the network propagator at all times. They are also pairwise commuting, being applied to disjoint Hilbert spaces. However, the cause should not be confused with the effect. A_{rs} projections shape or forge the unitary propagator with which they commute. To sum up, if the network admits one or more solutions, measurement at $t = \frac{\pi}{2\omega}$ gives one of them (that it is a solution is checkable in polynomial time). If the network admits no solution, conditions (i) through (iv) make up an impossible system. Measuring the network final state – at $t = \frac{\pi}{2\omega}$ – gives a non-solution. This is checkable in polynomial time and tells that the network is not satisfiable.

It is clear from the above that Connections “cut” network complexity, inducing a divide-and-conquer strategy. This diakoptic approach would make NP-complete \equiv P. However, we have been applying reverse engineering: until now the A_{rs} projections are just a nice-to-have feature. This raises the problem whether this feature can be physical.

5 Induced Symmetry

A_{rs} symmetry will be shown to be an epiphenomenon of fermionic antisymmetry in a special physical situation. This is generated by submitting a couple of identical fermions 1 and 2 to a suitable Hamiltonian^[12]. We assume that each fermion has two compatible, binary degrees of freedom χ and λ . Just for the sake of visualization (things should remain more abstract), we can think that each fermion is a spin 1/2 particle which can occupy one of either two sites of a spatial lattice. χ can thus become the particle spin component σ_z ($\chi = 0, 1$ correspond to $\sigma_z = \text{down}, \text{up}$) and $\lambda = r, s$ the label of the site occupied by the particle. For example, $|0\rangle_1 |1\rangle_2 |r\rangle_1 |s\rangle_2$ reads: σ_z of particle 1 down (0), σ_z of particle 2 up (1), site of particle 1 $\equiv r$, site of particle 2 $\equiv s$. There are 16 combinations like this, which make up the basis of the Hilbert space $H_{\lambda\chi}$. However, there are only six antisymmetrical combinations (not violating statistics) which make up the basis of the “symmetrical subspace” of $H_{\lambda\chi}$ (the

pun should not be misleading). These basis vectors are represented in first and second quantization and, when there is exactly one particle per site, in qubit notation (σ_z/λ stand for the qubit eigenvalue/label):

$$\begin{aligned} |a\rangle &= \frac{1}{\sqrt{2}} (|0\rangle_1 |1\rangle_2 - |1\rangle_1 |0\rangle_2) |r\rangle_1 |r\rangle_2 = a_{0r}^\dagger a_{1r}^\dagger |0\rangle, \\ |b\rangle &= \frac{1}{\sqrt{2}} (|0\rangle_1 |1\rangle_2 - |1\rangle_1 |0\rangle_2) |s\rangle_1 |s\rangle_2 = a_{0s}^\dagger a_{1s}^\dagger |0\rangle; \\ |c\rangle &= \frac{1}{\sqrt{2}} |0\rangle_1 |0\rangle_2 (|r\rangle_1 |s\rangle_2 - |s\rangle_1 |r\rangle_2) = a_{0r}^\dagger a_{0s}^\dagger |0\rangle = |0\rangle_r |0\rangle_s, \\ |d\rangle &= \frac{1}{\sqrt{2}} |1\rangle_1 |1\rangle_2 (|r\rangle_1 |s\rangle_2 - |s\rangle_1 |r\rangle_2) = a_{1r}^\dagger a_{1s}^\dagger |0\rangle = |1\rangle_r |1\rangle_s, \\ |e\rangle &= \frac{1}{2} (|0\rangle_1 |1\rangle_2 + |1\rangle_1 |0\rangle_2) (|r\rangle_1 |s\rangle_2 - |s\rangle_1 |r\rangle_2) = \\ &= \frac{1}{\sqrt{2}} (a_{0r}^\dagger a_{1s}^\dagger + a_{1r}^\dagger a_{0s}^\dagger) |0\rangle = \frac{1}{\sqrt{2}} (|0\rangle_r |1\rangle_s + |1\rangle_r |0\rangle_s), \\ |f\rangle &= \frac{1}{2} (|0\rangle_1 |1\rangle_2 - |1\rangle_1 |0\rangle_2) (|r\rangle_1 |s\rangle_2 + |s\rangle_1 |r\rangle_2) = \\ &= \frac{1}{\sqrt{2}} (a_{0r}^\dagger a_{1s}^\dagger - a_{1r}^\dagger a_{0s}^\dagger) |0\rangle = \frac{1}{\sqrt{2}} (|0\rangle_r |1\rangle_s - |1\rangle_r |0\rangle_s). \end{aligned}$$

$a_{\chi\lambda}^\dagger$ creates a particle of spin χ in site λ ; creation/annihilation operators form the algebra: $\{a_i^\dagger, a_j^\dagger\} = \{a_i, a_j\} = 0$, $\{a_i^\dagger, a_j\} = \delta_{i,j}$. Now we introduce the Hamiltonian $H_{rs} = E_a |a\rangle \langle a| + E_b |b\rangle \langle b| + E_c |c\rangle \langle c| + E_d |d\rangle \langle d|$ or, in second quantization $H_{rs} = -(E_a a_{0r}^\dagger a_{1r}^\dagger a_{0r} a_{1r} + E_b a_{0s}^\dagger a_{1s}^\dagger a_{0s} a_{1s} + E_c a_{0r}^\dagger a_{0s}^\dagger a_{0r} a_{0s} + E_d a_{1r}^\dagger a_{1s}^\dagger a_{1r} a_{1s})$, with $E_a, E_b, E_c, E_d \geq E$ discretely above 0. This leaves us with two degenerate ground eigenstates:

$$|e\rangle = \frac{1}{\sqrt{2}} (|0\rangle_r |1\rangle_s + |1\rangle_r |0\rangle_s) \text{ and } |f\rangle = \frac{1}{\sqrt{2}} (|0\rangle_r |1\rangle_s - |1\rangle_r |0\rangle_s).$$

Alternatively, their linear combinations $|0\rangle_r |1\rangle_s$ and $|1\rangle_r |0\rangle_s$ can be used as the two orthogonal ground eigenstates. The generic ground state is thus:

$$|\Psi\rangle = \alpha |0\rangle_r |1\rangle_s + \beta |1\rangle_r |0\rangle_s, \text{ with } |\alpha|^2 + |\beta|^2 = 1, \quad (5.1)$$

which satisfies A_{rs} symmetry. Let $A_{12} |\Psi\rangle = \frac{1}{2} (1 - P_{12})$ be the antisymmetrization projector. Due to the anticommutation relations: $A_{12} |0\rangle_r |1\rangle_s = |0\rangle_r |1\rangle_s$ and $A_{12} |1\rangle_r |0\rangle_s = |1\rangle_r |0\rangle_s$; also, $A_{12} |0\rangle_r |0\rangle_s = |0\rangle_r |0\rangle_s$ and $A_{12} |1\rangle_r |1\rangle_s = |1\rangle_r |1\rangle_s$, provided that $|0\rangle_r |0\rangle_s = |c\rangle$ and $|1\rangle_r |1\rangle_s = |d\rangle$ (see further below) and without forgetting that these are excited states.

The Connection can be implemented by suitably operating on the ground state (5.1). We assume that the initial, ‘‘symmetrical’’ state of the Connection is given by eq. (3.2): $|\Psi(0)\rangle = \cos \vartheta |0\rangle_r |1\rangle_s + \sin \vartheta |1\rangle_r |0\rangle_s$. Then transformation (3.1) $[\rho_r(t) = Q_r(\omega t) \rho_r(0) Q_r^\dagger(\omega t)]$ is applied to qubit r under continuous A_{rs} projection. Let $|\Psi(t)\rangle$ be a free normalized vector of $H_{\lambda\chi}$. The Connection state at time t is obtained by submitting $|\Psi(t)\rangle$ to the following mathematically simultaneous conditions,

for all t :

- i) $A_{12} |\Psi(t)\rangle = |\Psi(t)\rangle$;
- ii) $\rho_r(t) = Tr_s(|\Psi(t)\rangle \langle \Psi(t)|) = \cos^2(\vartheta + \varphi) |0\rangle_r \langle 0|_r + \sin^2(\vartheta + \varphi) |1\rangle_r \langle 1|_r$;

- iii) the distance between the vectors before and after projection is minimum as specified in Section 3;
- iv) the expected Connection energy, $\langle \xi(t) \rangle = \langle \Psi(t) | H_{rs} | \Psi(t) \rangle$, is minimum. Since this minimum will always be *zero*, time ordering is irrelevant.

It is readily seen that the solution of this system is still $|\Psi(t)\rangle$ of eq. (3.3):

$$|\Psi(t)\rangle = \cos(\vartheta + \varphi) |0\rangle_r |1\rangle_s + \sin(\vartheta + \varphi) |1\rangle_r |0\rangle_s.$$

Simultaneous satisfaction of (i), i.e. fermionic antisymmetry seen as projection, and (iv) (which is satisfied by $\langle \xi(t) \rangle = 0$) originates the projection constraint $A_{rs} |\Psi(t)\rangle = |\Psi(t)\rangle$, as is readily seen. Therefore, if $\langle \xi(t) \rangle = 0$, namely if the operation on qubit r is performed *adiabatically*, we obtain the Connection. Since this computation is *reversible*^[7,8], namely it does not dissipate free energy (the result of *driving and shaping* is a unitary evolution), in principle the operation can be adiabatic and $\langle \xi(t) \rangle$ can always be zero. This is of course an idealization: actually we are highlighting a *speculative*, possible way of dealing with NP-complete problems.

By the way we should note that the tensor products $|0\rangle_r |0\rangle_s$ and $|1\rangle_r |1\rangle_s$ that would be projected off since they violate A_{rs} symmetry (see the counterfactual reasoning of Section 2), are *not* the antisymmetrical excited states $|c\rangle$ and $|d\rangle$ satisfying A_{12} . They would be instead the symmetrical states of $H_{\lambda\chi}$:

$$\begin{aligned} |c\rangle' &= |0\rangle_r |0\rangle_s = \frac{1}{\sqrt{2}} |0\rangle_1 |0\rangle_2 (|r\rangle_1 |s\rangle_2 + |s\rangle_1 |r\rangle_2), \\ |d\rangle' &= |1\rangle_r |1\rangle_s = \frac{1}{\sqrt{2}} |1\rangle_1 |1\rangle_2 (|r\rangle_1 |s\rangle_2 + |s\rangle_1 |r\rangle_2). \end{aligned}$$

Let us consider the first equation. $|c\rangle'$ has the same qubit notation as $|c\rangle$ although $\langle c|c\rangle' = 0$. Let γ (γ') be the amplitude of $|c\rangle$ ($|c\rangle'$); α is the amplitude of $|0\rangle_r |1\rangle_s$. The first diagonal element of $\rho_r(t)$ would thus be $\|\alpha\|^2 + \|\gamma\|^2 + \|\gamma'\|^2$. The assumption is that $\rho_r(t)$ rotation does not raise $\|\gamma\|$ (initially zero) and consequently the energy, but $\|\gamma'\|$ which would be immediately killed by A_{12} projection, thus always remaining zero as it should be. The same can be said for $|d\rangle'$ and the second diagonal element of $\rho_r(t)$.

Let us address the problem of creating many Connections, namely an H_{rs} Hamiltonian per network wire r, s (fig. 2a). These H_{rs} operate on disjoint pairs of qubits. Viewed as A_{rs} projectors (which is the case when $\langle \xi(t) \rangle = 0$), they are pairwise commuting. Still in the idealized case of adiabatic operation, the Connections operate independently of each other.

6 Conclusion

The notion of applying a particle statistics symmetry (or projection) to divide the quantum whole into parts without clipping its richness – here computation

speed-up^[9,10,11, among others] — introduces an engineering (diakoptic) perspective in the design of quantum mechanisms. For the time being, this notion is developed at an abstract level. This paper is the *exploration* of a possible, alternative form of quantum computation. Finding model Hamiltonians which implement the Hermitean matrix of Section 5 could possibly be the next step.

The interpretation of particle statistics symmetry as projection on a predetermined subspace is best modeled in a two-way (advanced and retarded in time) propagation scheme^[12,13,14].

Part of this work was completed during the 1997 Elsag Bailey I.S.I. Foundation research meeting on quantum computation. Thanks are due to A. Ekert, D. Finkelstein, L. Levitin, S. Lloyd, C. Macchiavello and T. Toffoli for useful suggestions.

References

1. G. Kron, "The piecewise solution of large-scale systems", London Macdonald (1963).
2. G. Castagnoli and M. Rasetti, *Int. J. Theor. Phys.*, **32**, 2335 (1993).
3. A. Barenco, D. Deutsch, A. Ekert, R. Jozsa, *Phys. Rev. Lett.* **74**, 4083 (1995).
4. D.P. Di Vincenzo, *Phys. Rev. A* **50**, 1015 (1995).
5. A. Barenco, D. Deutsch, A. Ekert, *Proc. R. Soc. London A* **449**, 669 (1995).
6. S. Lloyd, *Phys.Rev.Lett.* **75**, 346 (1995).
7. C.H. Bennett, "Logical Reversibility of Computation" *IBM J. Res. Dev.* **6**, 525 (1979).
8. E. Fredkin and T. Toffoli, *Int. J. Theor. Phys.* **21**, 219 (1982).
9. D. Deutsch and R. Jozsa, *Proc. Roy. Soc. London A* **439**, 553 (1992).
10. D.R. Simon, *Proceedings of the 35th Annual Symposium on the Foundation of Computer Science*, Los Alamitos, CA, **116** (1994).
11. P.W. Shor, *Proceedings of the 35th Annual Symposium on the Foundation of Computer Science*, Los Alamitos, CA, **124** (1994).
12. G. Castagnoli, "Quantum Computation Based on Retarded and Advanced Propagation", Boston PhysComp 96. Available on the Web ([HTTP://xxx.lanl.gov](http://xxx.lanl.gov)): quant-ph/9706019, to be published in Physica D.
13. J.G. Cramer, *Rev. Mod. Phys.* **58**, 647 (1989).
14. G. Castagnoli, "Merging quantum computation and particle statistics", to be published in *Int. J. Theor. Phys.* 1998.01.

Practical Free-Space Quantum Cryptography

R.J. Hughes, W.T. Buttler, P.G. Kwiat, S.K. Lamoreaux, G.G. Luther,
G.L. Morgan, J.E. Nordholt, C.G. Peterson, and C.M. Simmons

University of California, Los Alamos National Laboratory
Los Alamos, New Mexico 87545, USA

hughes@lanl.gov

<http://p23.lanl.gov/Quantum/quantum.html>

Abstract. An experimental free-space quantum key distribution (QKD) system has been tested over an outdoor optical path of ~ 1 km under nighttime conditions at Los Alamos National Laboratory. This system employs the Bennett 92 protocol; here we give a brief overview of this protocol, and describe our experimental implementation of it. An analysis of the system efficiency is presented, as well as a description of our error detection protocol, which employs a two-dimensional parity check scheme. Finally, the susceptibility of this system to eavesdropping by various techniques is determined, and the effectiveness of privacy amplification procedures is discussed. Our conclusions are that free-space QKD is both effective and secure; possible applications include the rekeying of satellites in low earth orbit.

1 Introduction

Quantum cryptography was introduced in the mid-1980s [1] as a new method for generating the shared, secret random number sequences, known as cryptographic keys, that are used in crypto-systems to provide communications security. The appeal of quantum cryptography is that its security is based on laws of Nature, in contrast to existing methods of key distribution that derive their security from the perceived intractability of certain problems in number theory [2], or from the physical security of the distribution process.

Since the introduction of quantum cryptography, several groups have demonstrated quantum key distribution (QKD) over multi-kilometer distances of optical fiber [3]-[10], and recent advances have led to demonstrations of QKD over free-space indoor optical paths of 205 m [11], and outdoor optical paths of 75 m [12]. These demonstrations increase the utility of QKD by extending it to line-of-site laser communications systems. Indeed there are certain key distribution problems in this category for which free-space QKD would have definite practical advantages (for example, it is impractical to send a courier to a satellite). We are developing QKD for use over line-of-sight paths, and here we report our results of free-space quantum key generation over outdoor optical paths of up to 950 m under nighttime conditions.

2 Quantum-Key Distribution

The faithful transmission of polarized single photons through a turbulent medium (the atmosphere), receiving them with non-negligible probability and detecting them against a high ambient background, appear to be serious obstacles to free-space QKD. However, these obstacles can be overcome by exploiting sub-nanosecond timing techniques, narrow wavelength filters [13,14] spatial filtering [11], and adaptive optics [15]. To define the problem, we will require the generation of $\sim 1,000$ secret key bits between a ground station and a low-earth orbit satellite (~ 300 km altitude) in one overhead pass (duration ~ 8 minutes). In the following analysis we will assume that the QKD transmitter (Alice) is at the ground station and the receiver is on the satellite (Bob).

2.1 Free-Space Single Photon Detection and Transmission

The operational wavelength for free-space QKD should be chosen for both good atmospheric transmission properties and high detection efficiency. We have chosen to work at 772 nm where the atmospheric transmission from surface to space can be as high as 80% and single-photon detectors with efficiencies as high as 65% are commercially available (silicon avalanche photodiodes: APDs). Furthermore, at these optical wavelengths depolarizing effects of atmospheric turbulence are negligible as is the amount of Faraday rotation experienced on a surface to satellite path.

In order to detect a single QKD photon it is necessary to know when it will arrive. However, there will be variations in transmission time through the atmosphere owing to turbulence induced variations in refractive index, with time scales of the order of 0.01 – 0.1 s. Therefore, the photon arrival time can be communicated to the receiver by using a bright (multi-photon) precursor reference pulse, transmitted 100 ns (say) ahead of each QKD photon. The bright pulse and the “single photon” (produced by highly attenuating the pulsed output of a semiconductor laser) would each be of a ~ 100 -ps duration. (Note: the temporal length of the bright pulse is not as restricted as the temporal length of the dim-pulse; in fact, the bright pulse only needs to be short enough to allow the detection of the bright- and dim-pulses within the time allowed by the transmission rate. We also note that the atmosphere is only weakly dispersive.) The received bright pulse would then allow the receiver to set a 1-ns time window (say) within which to look for the QKD photon. This short time window would reduce background photon counts dramatically, and these can be reduced further using narrow filters at the wavelength of the QKD photons as well as spatial filtering. For example, 1-nm interference filters can be used and even narrower atomic vapor filters ($\sim 10^{-3}$ nm) are possible.

We now consider the rate at which QKD photons would be received at a satellite from a ground station transmitter. We will assume 20-cm diameter optics at both the transmitter and satellite receiver, leading to a ~ 1 -m diameter diffraction-limited spot size at the 300-km altitude satellite. However, there will be beam-wander owing to turbulence which can be as much as ~ 10 times the

diffraction limit (i.e., 10 arc-seconds of wander) so that the photon collection efficiency at the satellite is $\sim 10^{-4}$. Thus, with a laser pulse rate of 10 MHz, one photon-per-pulse on average and an atmospheric transmission of $\sim 80\%$, photons would arrive at the detector at a rate of ~ 1 kHz. Then, with a 65% detector efficiency and allowing for the 25% intrinsic efficiency of the quantum cryptography protocol, a key generation rate of ~ 150 Hz is feasible. With a beam tilt feedback system to keep the beam directed onto the satellite the key rate could be increased by a factor of 100. We must also consider the error rate.

We first consider errors arising from background photons arriving at the satellite. On a night time orbit with a full moon a typical radiance observed at the satellite at the transmission wavelength would be $\sim 1 \text{ mW m}^{-2} \text{ str}^{-1} \mu\text{m}^{-1}$ or $\sim 4 \times 10^{16} \text{ photons s}^{-1} \text{ m}^{-2} \text{ str}^{-1} \mu\text{m}^{-1}$. On a night with a new moon we take the background to be $\sim 10^{15} \text{ photons s}^{-1} \text{ m}^{-2} \text{ str}^{-1} \mu\text{m}^{-1}$. We will assume that the receiver “sees” a solid angle \sim five times the apparent size of the source (i.e., 5 arc-seconds) and that there is a 1-nm bandwidth interference filter placed in front of the detector, giving a background photon arrival rate of ~ 150 Hz (full moon); and ~ 4 Hz (new moon). With a 1-ns long time window on the detector, the probability of a background photon detection would be $\sim 10^{-7}$ (full moon); and $\sim 2.5 \times 10^{-9}$ (new moon) per 1-ns window. The single photon detector would only be triggered for precursor bright pulses that impinge on the satellite, giving approximately 120 detector triggers per arriving QKD photon, or 800 detector triggers per detected QKD photon. The bit error rate (BER) from background photons would therefore be $\sim 10^{-4}$ (full moon); and $\sim 2 \times 10^{-6}$ (new moon). Assuming a detector dark count rate of 50 Hz the BER will be dominated by background photons during full moon periods, and by detector noise during a new moon, with a BER $\sim 5 \times 10^{-5}$.

On daytime orbits the background radiance would be very much larger, $\sim 10^{22} \text{ photons s}^{-1} \text{ m}^{-2} \text{ str}^{-1} \mu\text{m}^{-1}$. Nevertheless, with an atomic vapor filter the rate of arrival of background photons would only be ~ 40 kHz (assuming a 10^{-3} nm filter width). The BER from this background would then be $\sim 2\%$.

From this simple analysis we see that QKD between a ground station and a low-earth orbit satellite should be possible on night time orbits and even in full daylight. During the several minutes that a satellite would be in view of the ground station there would be adequate time to generate tens of thousands of raw key bits, from which a shorter error-free key stream of several thousand bits would be produced after error correction and privacy amplification. A cryptographically useful quantity of key material could therefore be generated for this application.

2.2 The Bennett 92 Protocol

A QKD procedure starts with the sender, “Alice,” generating a secret random binary number sequence. For each bit in the sequence, Alice prepares and transmits a single photon to the recipient, “Bob,” who measures each arriving photon and attempts to identify the bit value Alice has transmitted. Alice’s photon state preparations and Bob’s measurements are chosen from sets of non-orthogonal

Table 1. Observation Probabilities

Alice's Bit Value	"0"	"0"	"1"	"1"
Bob Tests With	"1"	"0"	"1"	"0"
Observation Probability	$p=0$	$p=\frac{1}{2}$	$p=\frac{1}{2}$	$p=0$

possibilities. For example, using the B92 protocol [16] Alice agrees with Bob (through public discussion) that she will transmit a horizontal-polarized photon, $|h\rangle$, for each "0" in her sequence, and a right-circular-polarized photon, $|r\rangle$, for each "1" in her sequence. Bob agrees with Alice to randomly test the polarization of each arriving photon with vertical polarization, $|v\rangle$, to reveal "1s," or left-circular polarization, $|\ell\rangle$, to reveal "0s." In this scheme, Bob will never detect a photon for which he and Alice have used a preparation/measurement pair that corresponds to different bit values, such as $|h\rangle$ and $|v\rangle$, which happens for 50% of the bits in Alice's sequence. However, for the other 50% of Alice's bits the preparation and measurement protocols use non-orthogonal states, such as $|h\rangle$ and $|\ell\rangle$, resulting in a 50% detection probability for Bob, as shown in Table 1. Thus, by detecting single-photons Bob identifies a random 25% portion of the bits in Alice's random bit sequence, assuming a single-photon Fock state with no bit loss in transmission or reception. This 25% efficiency factor is the price that Alice and Bob must pay for secrecy.

Bob and Alice reconcile their common bits through a public discussion by revealing the locations, but not the bit values, in the sequence where Bob detected photons; Alice retains only those detected bits from her initial sequence. The resulting detected bit sequences comprise the raw key material from which a pure key is distilled using classical error detection techniques. The single-photon nature of the transmissions ensures that an eavesdropper, "Eve," can neither "tap" the key transmissions with a beam splitter (BS), owing to the indivisibility of a photon [17], nor copy them, owing to the quantum "no-cloning" theorem [18]. Furthermore, the non-orthogonal nature of the quantum states ensures that if Eve makes her own measurements she will be detected through the elevated error rate she causes by the irreversible "collapse of the wavefunction [19]."

2.3 Quantum-Key Transmitter: Alice

The QKD transmitter for our experiments (Fig. 1) consisted of a temperature-controlled single-mode (SM) fiber-pigtailed diode laser, a fiber to free-space launch system, a 2.5-nm bandwidth notch-filter, a variable optical attenuator, a polarizing beam splitter (PBS), a low-voltage Pockels cell, and a $27\times$ beam expander. The diode laser wavelength is temperature adjusted to 772 nm, and the laser is configured to emit a short, coherent pulse of approximately 1-ns length, containing $\sim 10^5$ photons.

A computer control system (Alice) starts the QKD protocol by pulsing the diode laser at a rate previously agreed upon between herself and the receiving

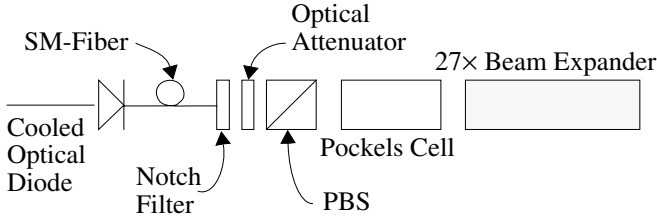


Fig. 1. QKD Transmitter.

computer control system (Bob). Each laser pulse is launched into free-space through the notch filter, and the ~ 1 ns optical pulse is then attenuated to an average of less than one photon per pulse, based on the assumption of a statistical Poisson distribution [20]. (The attenuated pulse only approximates a “single-photon” state; we tested the system with averages down to less than 0.1 photon per pulse. This corresponds to a 2-photon probability of $< 0.5\%$ and implies that less than 6 of every 100 detectable pulses will contain 2 or more photons, i.e., for a Poisson distribution, $P^{\bar{n}}$, with an average photon number of $\bar{n} = 0.1$, for every 1000 pulses there will be ~ 905 empty pulses, ~ 90 pulses of 1 photon, ~ 5 pulses of 2 photons, and ~ 1 pulse of 3 or more photons.) The photons that are transmitted by the optical attenuator are then polarized by the PBS, which transmits an average of less than one $|h\rangle$ photon to the Pockels cell. The Pockels cell is randomly switched to either pass the “single-photon” unchanged as $|h\rangle$ (zero-wave retardation) or change it to $|r\rangle$ (quarter-wave retardation). The random switch setting is determined by discriminating the voltage generated by a white noise source.

2.4 Quantum-Key Receiver: Bob

The free-space QKD receiver (Fig. 2) comprised a 8.9 cm Cassegrain telescope followed by the receiver optics and detectors. The receiver optics consisted of a 50/50 BS that randomly directs collected photons onto either of two distinct optical paths. The lower optical path contained a polarization controller (a quarter-wave retarder and a half-wave retarder), adjusted as an effective quarter-wave retarder, followed by a PBS to test collected photons for $|h\rangle$ (at first glance this may be confusing, but the effective quarter wave retarder converts $|h\rangle$ to $|r\rangle$ leading to a 50% probability an $|h\rangle$ photon will be detected); the upper optical path contained a half-wave retarder¹ followed by a PBS to test for $|r\rangle$ (again, perhaps

¹ A polarization controller was not required along the upper path because the 50/50 BS transmitted the P polarization (the component of polarization parallel to the plane of incidence) without introducing any phase shift, but the quarter- and half-wave retarder pair was necessary along the lower path because the BS reflected the P and S (component of polarization normal to the plane of incidence) polarizations differently, introducing some ellipticity to the reflected wave.

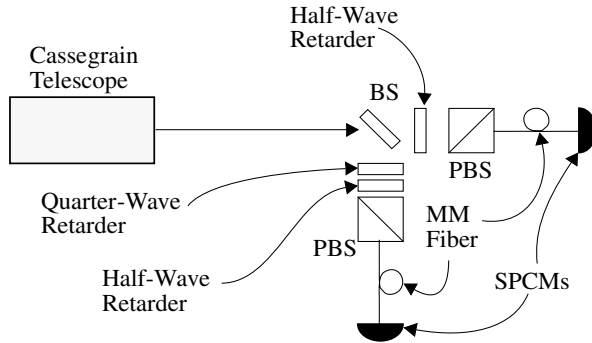


Fig. 2. QKD receiver.

confusing, but an $|r\rangle$ photon traveling this path is converted to $|\ell\rangle$, resulting in a 50% detection probability). The output port along each optical path was coupled by multi-mode (MM) fiber to a single-photon counting module (SPCM: EG&G part number: SPCM-AQ 142-FL). [Although the receiver did not include notch filters, the spatial filtering provided by the MM fibers effectively reduced noise caused by the ambient background during nighttime operations to negligible levels (the background was ~ 1.1 kHz).]

Bit values are determined in the following fashion: a single $|r\rangle$ photon traveling along the lower path encounters the polarization controller, and is converted to $|v\rangle$ and reflected away from the SPCM by the PBS, but a single $|h\rangle$ photon traveling the same path is converted to $|r\rangle$ and transmitted toward or reflected away from the SPCM in this path with equal probability; in contrast, a single $|h\rangle$ photon traveling the upper path is converted to $|v\rangle$ and reflected away from the SPCM in this path, but a single $|r\rangle$ photon traveling this path is converted to $|\ell\rangle$ and transmitted toward or reflected away from the SPCM with equal probability.

In this detection scheme, there are a total of four possible optical paths through the receiver, but only two of the paths, those which terminate upon the detectors seen in Fig. 2, contain definite polarization information (definite in the sense that Bob can know what polarization Alice has transmitted if one of these detectors fire). However, while the remaining two paths contain indeterminate polarization information (indeterminate in the sense that Bob cannot know with certainty whether Alice has transmitted $|h\rangle$, or $|r\rangle$ if a detector placed in either of these paths fires), but this information is important for the secure implementation of B92, as will be seen later (see Sec 4.1).

3 Outdoor Free-Space Experiments

The transmitter and receiver optics were operated over 240-, 500-, and 950-m outdoor optical paths, with the transmitter and receiver collocated in order to

simplify data acquisition. The various total optical path lengths were determined by positioning a 25.4 cm diameter mirror at the transmission distance half way point that reflected the transmitted beam back to the receiver. All measurements were made at night.

3.1 System Efficiency

In determining Bob's bit-rate, we consider that a BS partitions a weak photon stream in a binomial fashion [20]. We further assume that the effective wave retarders, combined with the PBSs, behave together as 50/50 BSs when analyzing non-orthogonal polarizations, i.e, if Alice transmits $|h\rangle$ and Bob analyzes with $|\ell\rangle$, or if Alice transmits $|r\rangle$ and Bob analyzes with $|v\rangle$. In addition, we treat the detectors as BSs with transmission coefficient $T_D = 0.65$, or in other words, that the detector, with efficiency $\eta_D = 0.65 = T_D$, also detects photon streams in a binomial way. We also treat the transmission and reception efficiency η —or power losses between the transmitter and receiver together with the losses which occur coupling power into the receiver's MM fibers—as random binomial processes. From the general form of the binomial probability distribution, we have

$$p_{\geq 1}^n = \sum_{m=1}^n \binom{n}{m} T^m R^{n-m}, \quad (1)$$

the probability that at least 1 photon out of n photons will be transmitted through the optical elements along the optical path (this is important because the detector responds to one or more photons). The net transmission probability is T , the reflection probability is R , and $T + R = 1$.

For calculation purposes, we use Eq. [2], which is equivalent to Eq. [1]

$$p_{\geq 1}^n \equiv 1 - (1 - \eta \cdot \eta_D \cdot 1/2 \cdot 1/2)^n, \quad (2)$$

where $T \mapsto \eta \cdot \eta_D / 4$, η and η_D are as previously defined, and the factor of $1/4 = 1/2 \cdot 1/2$ gives the probability that a photon collected at the receiver, of either $|h\rangle$, or $|r\rangle$, will be transmitted through the 50/50 BS followed by the effective quarter-wave retarder and PBS (for an $|h\rangle$ photon), or the half-wave retarder and PBS (for an $|r\rangle$ photon).

These binomial expanded products (Eq. [2]) of η , $1/4$, and η_D , are convolved with the Poisson probabilities, $P_n^{\bar{n}}$ that there will be exactly n photons in a pulse given that the average number of photons per pulse is \bar{n} :

$$P_n^{\bar{n}} = \frac{\bar{n}^n \exp(-\bar{n})}{n!}. \quad (3)$$

The convolution is then summed to give the detection probability as a function of the Poisson average photon number. This probability multiplied by the rate at which Alice transmits the coherent pulses, \mathcal{R}_A , gives the rate at which Bob detects 0s and 1s, \mathcal{R}_B :

$$\mathcal{R}_B = \mathcal{R}_A \sum_{n=1}^{\infty} P_n^{\bar{n}} [1 - (1 - \eta \eta_D / 4)^n]. \quad (4)$$

Table 2. A 200-Bit Sample of Alice’s (A) and Bob’s (B) Raw Key Material Generated by QKD over 1 km.

A	0000010101	1101101001	0000000000	0110010101
B	0000010101	1101101001	0000000000	0110010101
A	0011100010	0111011101	1110111000	0100100011
B	0011100010	0111011101	1110111000	0100100011
A	1110000000	0101101111	1001001010	0010000011
B	1110000000	0101101111	1001001010	0010000011
A	0000010111	0000111111	1111000000	1010101101
B	0000010111	0000111111	1101000000	1010101101
A	1111100111	1110111101	0100110100	1011101111
B	1111100011	1110111101	0100110100	1011101111

Our experimental result was $\mathcal{R}_B \sim 50$ Hz when the transmitter was pulsed at a rate of $\mathcal{R}_A = 20$ kHz, with $\bar{n} = 0.1$ photon per pulse for the 950-m path.

Finally, we note that in the limit that $\eta \cdot \eta_D \mapsto 1$, and given a Fock state of $m \equiv 1$ photon, then the photon probability distribution $P^n \mapsto \delta_{m-1}$, i.e., $\delta_{m-1} = 0 \forall m \neq 1$. In this limit—the limit of a perfect, lossless system—the sum vanishes and we are left with exactly 1 term $\mathcal{R}_B = \mathcal{R}_A/4$, which shows that Bob and Alice sacrifice 75% of their bits for privacy in agreement with Sec. 2.2

3.2 Error Rate

The bit error rate (BER) for the 950 m path was $\sim 1.5\%$ when the system was operating down to the < 0.1 photon per pulse level, where the BER is defined as the ratio of the bits received in error to the total number of bits received. A BER of $\sim 0.7\%$ was observed over the 240-m optical path and a BER of 1.5% was also observed over the 500 m optical path. A sample of raw key material from the 950-m experiment, with errors, is shown in Table 2

Spatial filtering reduced the ambient background (~ 1.1 kHz), and the narrow gated coincidence timing windows (~ 5 ns) reduced bit errors caused by the ambient background to less than ~ 1 every 9 s. Further, because detector dark noise (~ 80 Hz) contributed only about 1 dark count every 125 s, we believe that the BER was caused by misalignment and imperfections in the optical elements (wave-plates and Pockels cell).

3.3 Error Detection

Our experiments implement a two-dimensional (2D) parity check scheme that allows the generation of error-free key material. Error detection is accomplished by Bob and Alice organizing their reconciled bits (see Sec. 2.2) into 2D square matrices in the order that they were detected. Once organized, the parities of the rows and columns are determined and openly exchanged between Alice and Bob, and any column or row in which Bob and Alice possess different parities

is discarded. To ensure privacy, Alice and Bob also discard the bits oriented along the diagonals of their matrices. This guarantees the elimination of two bits for each row and column of the matrix, even when no errors are detected, eliminating knowledge revealed during the parity exchange.

Figure 3 illustrates the error detection protocol. In this example, Alice possesses the ‘good’ bits, and it is necessary for her and Bob to remove his ‘bad’ bits and distill error free key material. Bob possesses only two bad bits, but after openly communicating the column and row parities, they sacrifice good bits along the diagonals, and the 2 rows and 2 columns where parity differences were seen (parity differences are seen in columns 3 and 6 and rows 3 and 6). The net result, in this example, is 24 error-free bits: $key := \{100000110111110000010111\}$. Thus, in addition to the minimum 75% key lost during the B92 protocol, Bob and Alice have sacrificed another 62.5% of the detected bits. More complicated error detection codes could be employed to detect these as well, such as cyclic redundancy codes [22], but this was not done in our proof of principle experiment.

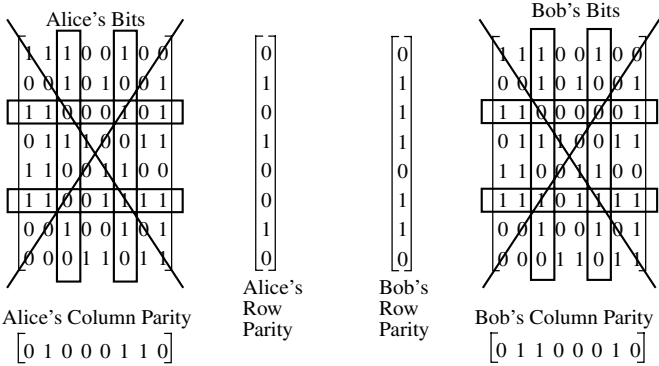


Fig. 3. Two-dimensional parity check scheme.

This is not the whole story, because the detection protocol does not detect all errors. For example, 2 errors in a column, combined with another error in a row containing one of the column errors (an ‘L’ shaped pattern), results in a missed bit-error. If there were 4 errors in a ‘box’ pattern, none of the errors would be detected, and so on.

We must emphasize, however, the strengths of the 2D routine as well. For example, the minimum Hamming distance [21][22], d , for a 2D scheme is the square of the minimum Hamming distance of the same detection scheme implemented in one-dimension (1D). (The Hamming distance tells how many errors can be detected, and/or corrected—one can detect $d - 1$ errors.) For our particular detection code, a parity check code, the minimum Hamming distance is 2 for

the 1D case, but in 2D this becomes 4. Once again, this is not the whole story, because there are situations in the 1D parity check scheme where more than one error can be detected, if the word is long enough; parity in 1D can detect an odd number of bit flips: 1, 3, 5, etc; however, even parity flips cannot be detected^[2]

To test our error detection scheme, we simulated random bit strings, with errors, and found that key material with bit errors as low as a few tenths of a percent had to be processed with the detection protocol at least twice to reduce errors to negligible levels. For random bit strings with high BERs (BERs of more than a couple of percent), small 2D matrices were needed on the first pass, but with each subsequent detection pass a larger 2D matrix could be used. We found that bit strings with BERs as high as 10% could be reduced to an estimated ~ 1 bit-error in a total of 10^9 bits after 4 passes, with $\sim 14\%$ of the initial key remaining; the sizes of the matrices in the 4 passes were 6 by 6, 7 by 7, 13 by 13, and 13 by 13, respectively.^[3] (We never operated our system with BERs this high, but in our simulations we wanted to determine the detection scheme's capabilities. We also found that there exists an optimal matrix size which most efficiently reduced errors while preserving a maximal amount of key material. The sizes varied from a 6 by 6 to a 12 by 12, almost linearly, for BERs between 10% and 1%.)

4 Eavesdropping: An Attack by Eve

Much has been said about the security of QKD against attack by an eavesdropper [\[19\]](#). There are essentially two types of attack to consider: opaque attacks and translucent attacks.

4.1 Opaque Attack

In an opaque attack, Eve intercepts all collectable bits, or single photons, by positioning herself between Alice and Bob. If Eve possesses a transmitter and receiver identical in every way to Bob's receiver and Alice's transmitter, and Bob, Alice and Eve are operating under the B92 protocol, then Eve can determine as much information about the key as could Bob. For example, if Alice's transmission basis is $|h\rangle$ and $|r\rangle$, and Eve's measurement basis is $|\ell\rangle$ and $|v\rangle$, then Eve can know Alice's transmitted bits with a maximum efficiency of 25%^[4]. If Eve retransmits the bits she "knows," then she will lower Bob's expected bit-rate, relative to Alice, by at least a factor of 4, but she will be forwarding bits of the correct value to Bob.

² We only detect and eliminate errors, and do not attempt to correct them.

³ We used square matrices of an odd size (7 by 7 and 13 by 13) in our simulations, but decided against using them in our experiments when we determined that the elimination of the diagonals of the matrices of odd size was insufficient to ensure security.

⁴ In a real system, Eve will experience reception losses associated with the collection, fiber launch and detection of the single photons.

If Eve can collect, measure, and quickly retransmit the bits she detects, she can then listen to Alice's and Bob's open bit reconciliation protocol (see Sec. 2.2). And, while Bob never reveals his bits values, Eve still knows what bits Alice and Bob commonly share because she knows when Alice began transmitted random bits. At this point, if Alice and Bob know their system well, Eve has been revealed by the additional factor of 4 attenuation, e.g., Eve has discarded a minimum of 75% of her bits, but Alice has discarded a minimum of 93.75% of her bits, i.e., Eve discards $(1 - 1/4)$, but Alice discards $(1 - 1/16)$.

Some could argue this additional attenuation to Bob's and Alice's common key is protection enough against an opaque attack, but our implementation of B92 adds another layer of protection if Eve attempts to bring Bob's bit rate to a rate indistinguishable from her own. Eve can do this by retransmitting a bright classical pulse to Bob for each single photon she detects.⁵ However, our system protects against this attack when operated in either a 2, 3, or 4 SPCM mode. In a 2 SPCM system, this type of attack would be revealed through an increase in "dual-fire" errors. Dual-fire errors occur when both SPCMs fire simultaneously. (In a perfect system there would be no dual-fire errors, regardless of the average photon number per pulse. However, in an imperfect experimental system dual-fire errors will occur, because there will be bit-errors associated with the transmission and measurement protocols, i.e., impure bit preparation and measurement associated with optical alignment of the transmission, receiving, and analysis optics.)

If we consider only a perfect system, then no matter how many horizontally polarized photons travel the $|r\rangle$ analysis path, none will reach the $|r\rangle$ analyzing detector. However, if this analysis path includes an effective half-wave retarder followed by a PBS, then the half wave-retarder will convert right-circular polarized photons to left-circular polarized photons which will then be equally split equally between the two output paths. If both paths are each followed by an SPCM then both SPCMs will fire.

The component of a right-circular polarized pulse that travels the $|h\rangle$ analysis path encounters the effective quarter-wave retarder followed by another PBS. The quarter-wave retarder converts this right-circular polarized 'bright' pulse to a vertical-polarized 'bright' pulse which is reflected along the path away from the $|h\rangle$ analyzing detector. If this path contains an SPCM, then this SPCM will fire together with the two SPCMs which terminate on the $|r\rangle$ analyzing path. Thus, 3 of 4 detectors have fired alerting Bob and Alice that Eve is opaquely attacking the key. A similar argument applies if Bob is using 3 detectors.

4.2 Translucent Attack

Eve could also passively, or translucently, attack the quantum transmission with a BS. In this scheme, Eve receives the binomial reflection probability of the BS she uses to reflect photons toward her receiving optics, and Bob receives the

⁵ In B92 it is possible to send bright classical pulses of the appropriate polarization to ensure that every bit transmitted is detected at the receiver.

binomial transmission probability of the BS Eve uses. In a translucent attack it is necessary to consider Eve's and Bob's reception and detection efficiencies, which are independent; Eq. 5 shows the amount of information on the key Eve receives as a function of the reflection coefficient, $R_E = 1 - T_E$, of BS she uses in her translucent attack, her receiver efficiency η_E , and her detector efficiency, η_D^E . Equation 6 shows the amount of key Bob receives as a function of the transmission coefficient, T_E , of the BS Eve uses in a translucent attack, and his reception and detection efficiencies, η_B and η_D^B .

$$\mathcal{R}_E = \mathcal{R}_A \sum_{n=1}^{\infty} P_n^{\bar{n}} \left[1 - \left(1 - \frac{\eta_E \eta_D^E (1 - T_E)}{4} \right)^n \right], \quad (5)$$

and

$$\mathcal{R}_B = \mathcal{R}_A \sum_{n=1}^{\infty} P_n^{\bar{n}} \left[1 - \left(1 - \frac{\eta_B \eta_D^B T_E}{4} \right)^n \right], \quad (6)$$

The 1/4 reduction of these products is as previously described in Sec 3. Eve's bit rate is \mathcal{R}_E , and \mathcal{R}_B is Bob's, and \mathcal{R}_A is the rate Alice is transmitting.

The privacy P , or the percentage of information Eve possesses on Alice's and Bob's common key, is determined as the ratio of the number of bits Eve and Bob share (observe coincidentally) to the number of bits Alice and Bob share. First of all, if there is only 1 photon in a pulse, then either Eve or Bob will receive it, but not both. Based on this premise, Eq. 7 shows the number of bits that Bob and Eve will share if Eve attacks the key with a BS of transmission coefficient T_E , and reflection coefficient $R_E = 1 - T_E$.

$$N_{B \wedge E} = \mathcal{R}_A \sum_{n=2}^{\infty} P_n^{\bar{n}} \sum_{m=1}^{n-1} \binom{n}{m} T_E^m R_E^{n-m} \left[1 - \left(1 - \frac{\eta_B \eta_D^B}{4} \right)^m \right] \left[1 - \left(1 - \frac{\eta_E \eta_D^E}{4} \right)^{n-m} \right] \quad (7)$$

Equation 8 shows Alice's and Bob's privacy, P .

$$P = \frac{\sum_{n=2}^{\infty} P_n^{\bar{n}} \sum_{m=1}^{n-1} \binom{n}{m} T_E^m R_E^{n-m} \left[1 - \left(1 - \frac{\eta_B \eta_D^B}{4} \right)^m \right] \left[1 - \left(1 - \frac{\eta_E \eta_D^E}{4} \right)^{n-m} \right]}{\sum_{n=1}^{\infty} P_n^{\bar{n}} \left[1 - \left(1 - \frac{\eta_B \eta_D^B T_E}{4} \right)^n \right]} \quad (8)$$

Under this type of translucent attack, if Eve uses 50/50 BS, and if Alice transmits coherent Poisson pulses with an average of 0.1 photon per pulse, and if Bob's and Eve's system and detection efficiencies are equal, then for every 250 bits Eve and Bob acquire, Eve will commonly share ~ 3 of her 250 bits with Bob's 250 bits, or $\sim 3/250$ of Alice and Bob's common key. In fact, because Eve's knowledge on Alice's and Bob's common key is coupled to hers and Bob's system efficiencies, this situation represents the maximum amount of information Eve can obtain on Alice's and Bob's common key even if her system is perfectly efficient and Bob's is not. The inverse is also true, i.e., if Bob's system is more

efficient than Eve's, then the amount of information Eve can determine on Alice's and Bob's common key decreases. (Note: at this point, we have only been able to show this empirically, but we have found no exceptions to these facts). Eve could determine which bits she commonly shares with Bob when Alice and Bob reconcile their common bits.

Finally we note that because Alice transmits coherent states, as opposed to single photon Fock states, she and Bob also need to add a stage of "privacy amplification [23]" to reduce any partial knowledge gained by an eavesdropper to less than 1-bit of information. We have not implemented such a privacy amplification protocol at this time, but our free-space QKD system does incorporate "one time pad [24]" encryption—also known as the Vernam Cipher: the only provably secure encryption method—and could also support any other symmetric key system.

5 Conclusions

The results in this paper demonstrate free-space QKD through a turbulent medium under nighttime conditions. We have described a system that provides two parties a secure method to secretly communicate with a simple system based on the B92 protocol. We presented two attacks on this protocol and demonstrated the protocol's built in protections against them. This system was operated at a variety of average photon number per pulse down to an average of < 0.1 photon per pulse. The results were achieved with low BERs, and the 240-m experiment demonstrated that BERs of 0.7% or less are achievable with this system. This protocol could be implemented with classical signature authentication [2] and privacy amplification procedures to ensure the security of private information. From these results we believe that it will be feasible to use free-space QKD for re-keying satellites in low-earth orbit from a ground station.

References

1. C. H. Bennett, and Brassard, G.: Quantum cryptography: public key distribution and coin tossing. Proc. of IEEE Int. Conf. on Comp., Sys., and Sig. Proc., Bangalore, India (1984) 175.
2. A. J. Menezes, van Oorschot, P. C., and Vanstone, S. A.: *Handbook of Applied Cryptography*. CRC Press, New York (1997).
3. A. Muller, Breguet, J., and Gisin, N.: Experimental demonstration of quantum cryptography using polarized photons in optical fibre over more than 1 km. Europhys. Lett. **23** (1993) 383.
4. A. Muller, Zbinden, H., and Gisin, N.: Quantum cryptography over 23 km in installed under-lake telecom fiber. Europhys. Lett. **33** (1996) 335.
5. P. D. Townsend, Rarity, J. G., and Tapster, P. R.: Enhanced single-photon fringe visibility in a 10 km-long prototype quantum cryptography channel. Elec. Lett. **29** (1993) 634.
6. C. Marand, and Townsend, P. D.: Quantum key distribution over distances as long as 30 km. Opt. Lett. **20** (1995) 1695.

7. J. D. Franson, and Ilves, H.: Quantum cryptography using optical fibers. *Appl. Opt.* **33** (1994) 2949.
8. R. J. Hughes, Alde, D. M., Dyer, P., Luther, G. G., Morgan, G. L., and Schauer, M.: Quantum cryptography. *Contemp. Phys.* **36** (1995) 149.
9. R. J. Hughes, Luther, G. G., Morgan, G. L., Peterson, C. G., and Simmons, C.: Quantum cryptography over underground optical fibers. *Lecture Notes In Computer Science* **1109** (1996) 329.
10. R. J. Hughes, Buttler, W. T., Kwiat, P. G., Luther, G. G., Morgan, G. L., Nordholt, J. E., Peterson, C. G., and Simmons, C. M.: Secure communications using quantum cryptography. *Proc. of SPIE* **3076** (1997) 2.
11. W. T. Buttler, Hughes, R. J., Kwiat, P. G., Luther, G. G., Morgan, G. L., Nordholt, J. E., Peterson, C. G., and Simmons, C. M.: Free-space quantum-key distribution. Scheduled for *Phys. Rev. A* **57** (1998).
12. B. C. Jacobs, and Franson, J. D.: Quantum cryptography in free space. *Opt. Lett.* **21** (1996) 1854.
13. J. G. Walker, Seward, S. F., Rarity, J. G., and Tapster, P. R.: Range measurement photon by photon. *Quant. Opt.* **1** (1989) 75.
14. S. F. Seward, Tapster, P. R., Walker, J. G., and Rarity, J. G.: Daylight demonstration of a low-light-level communication system using correlated photon pairs. *Quant. Opt.* **3** (1991) 201.
15. C. A. Primmerman, Murphy, D. V., Page, D. A., Zollars, B. G., and Barclay, H. T.: Compensation of atmospheric optical distortion using a synthetic beacon. *Nature (London)* **353** (1991) 141.
16. C. H. Bennett: Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.* **68** (1992) 3121.
17. J. F. Clauser: Experimental distinction between quantum and classical field-theoretic predictions for photoelectric effect. *Phys. Rev. D* **9** (1974) 853.
18. W. K. Wothers, and Zurek, W. H.: A single quantum cannot be cloned. *Nature (London)* **299** (1982) 802.
19. A. K. Ekert, Huttner, B., Palma, G. M., and Peres, A.: Eavesdropping on quantum cryptosystems. *Phys. Rev. A* **50** (1994) 1047.
20. B. E. A. Saleh, and Teich, M. C.: *Fundamentals of Photonics*. Ch. 11, Jon Wiley and Sons, Inc., New York (1991).
21. R. W. Hamming: *Coding and Information Theory*. Prentice Hall, New Jersey (1980).
22. J. Wakerly: *Error Detecting Codes, Self-Checking Circuits and Applications*. North-Holland, New York (1978).
23. C. H. Bennett, Brassard, G., Crepeau, C., and Maurer, U. M.: Generalized privacy amplification. *IEEE Trans. Inf. Th.* **41** 1915 (1995).
24. G. S. Vernam: Cipher printing telegraph systems for secret wire and radio telegraphic communications. *Trans. Am. Inst. Electr. Eng.* **XLV** (1926) 295.

Quantum Cryptography, Eavesdropping, and Unsharp Spin Measurement

S. Roy and G. Kar

Physics and Applied Mathematics Unit
Indian Statistical Institute
203, B. T. Road
Calcutta - 700 035
`sisir@isical.ernet.in`

Abstract. The various eavesdropping strategies on a quantum cryptographic channel as discussed by Gisin et.al. can be reproduced by unsharp spin (represented by positive operator valued measure) measurement of spin-1/2 particle. An upper bound for unsharp parameter is obtained so that it still gives rise to violation of Bell's inequality for the two particle transmission channel.

Recently Gisin and Huttner [1] analysed various eavesdropping strategies on a quantum cryptographic channel. In their treatment with a 2-D probe, they have taken a particular evolution and have shown that for spin-1/2 particle, Eve's eavesdropping simply shrink the Bloch vector of the Alice's state by a factor η which is a function of measurement of intensity γ ($0 \leq \gamma \leq \pi/2$). Then they have applied the same kind of interaction in the case of cryptographic protocol with two spin-1/2 particles in a singlet state and have calculated the parameter S that appears in the CHSH version of Bell's inequality as a function of the same γ .

In this paper, we will reproduce these particular results applying unsharp spin measurement formalism [2] which represents Eve's interference with the particle reaching to Bob.

For this purpose let us shortly describe the unsharp measurement of spin property. In unsharp formalism the spin property is represented by POV measure which is more general than projection operator. We write the POV measure representing the unsharp property [2,3]

$$E_\lambda(n) = \frac{1}{2}[I + \lambda n \cdot \sigma] \quad (1)$$

where n is a unit vector and λ is real with $0 \leq \lambda \leq 1$. For $\lambda = 1$, it becomes sharp spin property $E(n)$.

$E_\lambda(n)$ can be written as

$$E_\lambda(n) = \frac{1+\lambda}{2}E(n) + \frac{1-\lambda}{2}E(-n) \quad (2)$$

where

$$E(-n) = (I - E(n))$$

$\frac{1+\lambda}{2} = r > 1/2$ is called degree of reality and $\frac{1-\lambda}{2} = u < 1/2$ is called degree of unsharpness for the spin property.

Now every measurement corresponds to to an operation [4] which gives the final state of the system. Here we consider the generalised Luder operation [4,5] which disturbs the initial state minimally. The Luder generalised operation ϕ_L corresponding to the measurement $E_\lambda(n)$ on a state represented by the density operator $\rho = \frac{1}{2}[I + a.\sigma]$ is given by

$$\begin{aligned} \phi_L \rho &= (E_\lambda(n))^{1/2} \rho (E_\lambda(n))^{1/2} \\ &= \frac{1}{2} \text{Tr}[\rho E_\lambda(n)] + \frac{1}{4} \{ (1 - \lambda^2)^{1/2} [a - n(n.a)] + n[\lambda + (a.n)] \} . \sigma \end{aligned} \quad (3)$$

Let the state prepared by Alice is given by

$$\rho_{Alice} = \frac{1}{2}[I + m.\sigma] \quad (4)$$

where m is an unit vector known as Bloch vector representing ρ_{Alice} on the Poincare sphere.

Alice sends this state to Bob. But before reaching to Bob, Eve makes some unsharp measurement on this state. So the changed state due to the measurement disturbance (without reading the result) which reaches to Bob is given by

$$\rho_{Bob} = (E_\lambda(n))^{1/2} \rho_{Alice} (E_\lambda(n))^{1/2} + (I - E_\lambda(n))^{1/2} \rho_{Alice} (I - E_\lambda(n))^{1/2} \quad (5)$$

Applying (3) we get

$$\rho_{Bob} = \frac{1}{2}I + \frac{1}{2}[m + (1 - \lambda^2)^{1/2}\{(n.m)n - m\}].\sigma \quad (6)$$

But from such state reaching to Bob, Alice and Bob can easily infer that the noise in that the transmission is not produced by random process and hence get some information about Eve's strategy. For avoiding that Eve must perform two spin measurements, one along direction n and another along \bar{n} (along perpendicular direction to n) randomly.

From $E_\lambda(\bar{n})$ measurement, the resulting state ρ'_{Bob} will be given by

$$\rho'_{Bob} = \frac{1}{2}I + \frac{1}{2}[m + (1 - \lambda^2)^{1/2}\{(\bar{n}.m)\bar{n} - m\}].\sigma \quad (7)$$

So the final density matrix for the random combination of both the strategies can be written as

$$\bar{\rho}_{Bob} = \frac{1}{2}\rho_{Bob} + \frac{1}{2}\rho'_{Bob} = \frac{1}{2}[I + \frac{1}{2}\{1 + (1 - \lambda^2)^{1/2}\}m.\sigma] \quad (8)$$

which is same as the density matrix in equation (10) of reference 1 with

$$(1 - \lambda^2)^{1/2} = \cos \gamma$$

The effect of Eve's eavesdropping is thus simply to shrink the Bloch vector by the factor

$$\frac{1}{2}\{1 + (1 - \lambda^2)^{1/2}\}$$

which is a function of the unsharp parameter λ .

Let us now turn to the strategy of cryptography with the EPR pair one with Alice and the other reaching to Bob. On the later Eve makes some unsharp spin measurement. We want to see how much unsharpness is allowed for spin measurement so that Alice and Bob still get the violation of Bell's inequality.

We write the singlet state as

$$\psi_0 = \frac{1}{\sqrt{2}}[\psi_n^1 \otimes \psi_{-n}^2 - \psi_{-n}^1 \otimes \psi_n^2] \quad (9)$$

where particle 1 reaches to Alice and 2 reaches to Bob and ψ_{\pm} are eigen states of σ_z . The state after Eve's unsharp spin measurement on the particle 2 will be given by

$$\begin{aligned} W &= I \otimes (E_{\lambda}(n))^{1/2} P[\psi_0] I \otimes (E_{\lambda}(n))^{1/2} \\ &+ | \otimes (I - E_{\lambda}(n))^{1/2} P[\psi_0] I \otimes (I - E_{\lambda}(n))^{1/2} \\ &= \frac{1}{2} P[\psi_n^1 \otimes \psi_{-n}^2] + \frac{1}{2} P[\psi_{-n}^1 \otimes \psi_n^2] \\ &- \frac{1}{2} |\psi_n^1 \rangle \langle \psi_{-n}^1| \otimes |\psi_{-n}^2 \rangle \langle \psi_n^2| - c.c \end{aligned} \quad (10)$$

Here $P[.]$ represents projection operator on the vector in the square bracket.

Now the expectation value of the Bell operator (for the choice of spin observables giving maximal violation in a singlet state) for the state W is

$$\langle B_{Bell} \rangle_W = \sqrt{2}[1 + (1 - \lambda^2)^{1/2}] \quad (11)$$

Again Bell's inequality satisfies

$$-2 \leq \langle B_{Bell} \rangle \leq 2$$

So for violation of Bell's inequality we need

$$\lambda < \sqrt{2}(\sqrt{2} - 1)^{1/2} \quad (12)$$

In conclusion, it is to be cleared that we completely agree to the result obtained by Gisin et.al. for various Eavesdropping strategies. We only showed that their particular measurement scheme for producing the results as discussed above, is equivalent to unsharp spin measurement with generalised Luder operation. In cryptography, this equivalency is important because sometimes the unsharp measurement produce the same information as its sharp counterpart but disturbs the initial state minimally.

REFERENCES

- [1] N. Gisin, B, Huttner, Phys. Lett. A. **228** (1997)13.
- [2] P. Busch, Found. Phys. **17** (1987)905.
- [3] G. Kar, Int. J. Theot. Phys. **35** (1996)1279.
- [4] K. Kraus, States, effects and operations, Springer-Verlag, 1983.
- [5] P. Busch, Phys. Rev. D. **33** (1986)2253.

Information-Theoretic Aspects of Quantum Copying

Nicolas J. Cerf*

W. K. Kellogg Radiation Laboratory, California Institute of Technology,
Pasadena, California 91125
Information and Computing Technologies Research Section, Jet Propulsion
Laboratory, Pasadena, California 91109
Center for Nonlinear Phenomena and Complex Systems, Université Libre de
Bruxelles, 1050 Bruxelles, Belgium

Abstract. An information-theoretic approach to quantum copying is discussed, relying on the notion of quantum loss, a quantity that reflects the transmission quality in a noisy quantum channel. More specifically, an entropic *no-cloning inequality* is derived for a Hilbert space of arbitrary dimension, which describes the tradeoff between the losses of the channels leading to the two copies. Then, focusing on quantum bits, a family of *Pauli cloning machines* is introduced. These machines produce two imperfect copies of a single quantum bit that emerge from two distinct Pauli channels. The balance between the quality of the two copies is shown to result from a genuine *complementarity principle*. In the special case where the two outputs are associated with depolarizing channels of probability p and p' , the domain in $(\sqrt{p}, \sqrt{p'})$ -space located inside a particular ellipse representing close-to-perfect cloning is forbidden. Finally, the class of symmetric Pauli cloning machines is used to provide an upper bound on the quantum capacity of the Pauli channel of probabilities p_x , p_y and p_z . The capacity is proven to be vanishing if $(\sqrt{p_x}, \sqrt{p_y}, \sqrt{p_z})$ lies outside an ellipsoid whose pole coincides with the depolarizing channel that underlies the universal cloning machine.

1 Introduction

A remarkable property of quantum information is that it cannot be copied, in contrast with information we are used to in classical physics. This means that there exists no physical process that can produce *perfect* copies of a system that is initially in an *unknown* quantum state. This so-called *no-cloning* theorem, recognized by Dieks [1] and Wootters and Zurek [2], is an immediate consequence of the linearity of quantum mechanics, and lies at the heart of quantum theory. Indeed, if perfect cloning *was* permitted, the Heisenberg uncertainty principle could be violated by measuring conjugate observables on many copies of a single quantum system.

* This paper was presented at the 1st NASA International Conference on Quantum Computing and Quantum Communications, Palm Springs, February 1998.

Consider a cloning machine that duplicates a quantum bit (a 2-state system) that is initially in an arbitrary state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ whose amplitudes are unknown. It is easy to build such a machine that perfectly copies the two basis states $|0\rangle$ and $|1\rangle$, but then it badly duplicates the superpositions $2^{-1/2}(|0\rangle \pm |1\rangle)$. In other words, it cannot produce *perfect* copies of *all* possible input states. This being so, we may ask how well one can *approximately* duplicate the unknown state of a quantum bit (qubit) if the quality of the copies is required to be independent of the input state. This question was answered by Buzek and Hillery who first showed that it is possible to construct a cloning machine that yields two *imperfect* copies of a single qubit [3]. Specifically, a universal cloning machine (UCM) can be defined that creates two copies characterized each by the same density operator ρ , the fidelity of cloning being $f \equiv \langle\psi|\rho|\psi\rangle = 5/6$. This machine is called *universal* because it produces copies that are *state-independent*: both output qubits emerge from a depolarizing channel of probability $1/4$, that is, the Bloch vector characterizing the input qubit is shrunk by a factor $2/3$ regardless its orientation. The UCM was later proved to be optimal by Bruss et al. [4], and Gisin and Massar [5]. Much attention has been recently directed towards quantum cloning machines, because of their use in connection with quantum communication and cryptography (see, e.g., [4,6]).

The outline of the paper is as follows. In Section 2, we start by reviewing the characterization of a noisy quantum channel by its *loss*, a quantity that reflects the quality of the transmission. The loss, depending on the input and the operation performed by the channel, can be shown to vanish when the transmission of quantum information is perfect. This concept is used to display the information-theoretic significance of the quantum no-cloning theorem. More precisely, an *entropic no-cloning inequality* is derived, characterizing the impossibility of copying imposed by quantum mechanics: $L_A + L_B \geq 2S$, where L_A and L_B are the losses characterizing outputs A and B , respectively, while S is the source entropy. In Section 3, we introduce a family of asymmetric *Pauli cloning machines* (PCM), which produces two *distinct* (approximate) copies of a single quantum bit, each emerging from a Pauli channel [7]. This is in contrast with the cloning machines considered in the literature, which are symmetric (both outputs being characterized by the same density operator). The family of PCMs relies on a parametrization of 4-qubit wave functions for which all qubit pairs are in a mixture of Bell states.

Using a particular class of asymmetric PCMs whose outputs emerge from (distinct) depolarizing channels, we derive a *no-cloning uncertainty relation* governing the tradeoff between the quality of the copies of a quantum bit: $a^2 + ab + b^2 \geq 1$, where a^2 and b^2 are the depolarizing fractions of the channels associated with outputs A and B , respectively. It is, by construction, a tight inequality which is saturated using our PCM. More generally, the complementarity between the two copies produced by a Pauli cloning machine is shown to result from an *uncertainty principle*, much like that associated with Fourier transforms. This uncertainty principle relates the probability distributions underlying the channels leading to the two outputs of the cloner. Finally,

the subclass of *symmetric* PCMs is used in order to express an upper bound on the quantum capacity of the Pauli channel. In particular, the capacity of the Pauli channel of probabilities $p_x = x^2$, $p_y = y^2$ and $p_z = z^2$, is shown to vanish if (x, y, z) lies outside the ellipsoid $x^2 + y^2 + z^2 + xy + xz + yz = 1/2$, whose pole coincides with the depolarizing channel underlying the UCM. This implies an upper bound on the capacity C of any Pauli channel associated with a point (x, y, z) located inside the ellipsoid, namely $C \leq 1 - 2(x^2 + y^2 + z^2 + xy + xz + yz)$.

2 Information-Theoretic Significance of the Quantum No-cloning Theorem

2.1 Entropic Characterization of a Noisy Quantum Channel

Let us outline the entropic treatment of a noisy quantum channel that is introduced in Refs. [8,9]. The description of a channel involves three quantum systems of arbitrary dimensions: X (the input quantum system whose processing by the channel is concerned), R (a reference system which X is initially entangled with), and E (an environment which X is interacting with in the noisy channel, inducing decoherence) [10]. More specifically, we assume that X is initially entangled with R , so that the joint state of X and R is the *pure* state $|\Psi_{RX}\rangle$. We may as well regard X as a quantum source, being initially in a mixed state ρ_X (realized by a given ensemble of quantum states associated with some probability distribution). The “purification” of ρ_X into $|\Psi_{RX}\rangle$ can always be achieved by extending the Hilbert space \mathcal{H}_X to \mathcal{H}_{RX} , so that we have $\rho_X = \text{Tr}_R(|\Psi_{RX}\rangle\langle\Psi_{RX}|)$. The corresponding reduced von Neumann entropies are

$$S(R) = S(X) \equiv S \quad (1)$$

where $S = -\text{Tr}_X(\rho_X \log \rho_X)$ is called the *source entropy* and is a function of ρ_X , the density operator characterizing the input X . In the dual picture where an *arbitrary* state of X (rather than entanglement) is sent through the channel, S then measures the “arbitrariness” of the input X (it can be viewed as the average number of quantum bits that must be processed by the channel in order to transmit the state of X). In our information-theoretic characterization of quantum channels, we prefer to consider an input X that is entangled with R , so that we investigate the transmission of entanglement rather than of arbitrary states. The initial quantum *mutual entropy* to be transmitted in a quantum channel is thus

$$S(R:X) = S(R) + S(X) - S(RX) = 2S \quad (2)$$

that is, twice the source entropy.

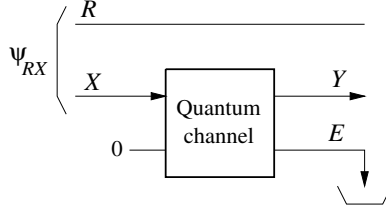
When it is processed by the noisy quantum channel, X interacts with E (assumed to be initially in a pure state $|0\rangle$) according to a particular unitary transformation U , inducing decoherence (see Fig. [1]). This depicts the most general (trace-preserving) operation of a quantum channel that is allowed by quantum mechanics. Roughly speaking, only a fraction of the initial entanglement with R

can be recovered in general after processing by the channel, while the rest of the entanglement with R is lost (i.e., it is transferred to the environment E). More specifically, the quantum output Y (i.e., the decohered quantum system after interaction with E) is characterized by the density operator

$$\rho_Y = \text{Tr}_E(U(\rho_X \otimes |0\rangle\langle 0|)U^\dagger) \quad (3)$$

The completely positive linear map $\rho_X \rightarrow \rho_Y$ corresponds to the “quantum operation” performed by the noisy channel [10].

Fig. 1. Noisy quantum channel of input X (initially entangled with a reference R) and output Y . Decoherence is induced by the interaction with an environment E , initially in the pure state $|0\rangle$.



The processing of quantum information through the channel $X \rightarrow Y$ can be characterized by two entropies, the quantum mutual entropy I and the quantum loss L :

$$I = S(R:Y) \quad (4)$$

$$L = S(R:E) \quad (5)$$

where E denotes the environment *after* decoherence while R is the reference before or after decoherence (R is not involved in decoherence). It can be shown that I and L are independent of the choice of the reference system R provided that the latter purifies the input X . They depend in general on the channel input (i.e., ρ_X) and on the quantum operation performed by the channel (i.e., the completely positive trace-preserving map $\rho_X \rightarrow \rho_Y$ that is specified by U in the joint space of X and E). The processing of quantum information in the noisy channel is characterized by the balance between I and L , these two quantities always summing to *twice* the source entropy:

$$I + L = 2S \quad (6)$$

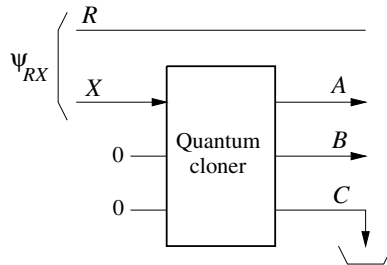
Physically, I represents the amount of the initial mutual entropy with respect to R (i.e., $2S$) that has been processed by the channel, while L corresponds to the fraction of it that has been unavoidably lost in the environment. If the channel is *lossless* ($L = 0$), then a perfect transmission of quantum information can be achieved by applying an appropriate decoding [8,9,10]. In other words,

the interaction with the environment can be perfectly “undone”, and the initial entanglement with R can be fully recovered. (Equivalently, this means that an arbitrary initial state can be recovered without error.) Conversely, if $I = 0$, then no information (classical *or* quantum) can be processed by the channel. In between these limiting cases, the loss provides a measure of (or rather a bound on) how reliably quantum information can be transmitted [9].

2.2 Entropic No-cloning Inequality

Let us consider a cloning machine of input X and outputs A and B . In other words, A and B are the (approximate) copies of X produced by the cloning machine. The above entropic analysis can be used by associating the input-to-output overall operation characterizing each copy with a noisy quantum channel. For that purpose, we assume as before that the input of the cloner, X , is initially entangled with a reference system R . The question will be to determine to what extent entanglement (with respect to R) can be transmitted simultaneously to both outputs, A and B . To be general, we also assume that the cloning machine, denoted as C , is itself involved in the copying process. Thus, the action of the cloner can be described, in full generality, as a particular unitary operation acting on X together with two auxiliary systems (each being initially in a prescribed state $|0\rangle$), yielding A , B , and C (see Fig. 2).

Fig. 2. Quantum cloning machine of input X (initially entangled with a reference R) and outputs A and B . The symbol C refers to an ancilla or the cloning machine.



Thus, the first output A emerges from the channel $X \rightarrow A$, which can be characterized by the loss

$$L_A = S(R:BC) \quad (7)$$

since B and C play the role of an environment for this channel. Similarly, the second output B emerges from the channel $X \rightarrow B$, which is associated with a loss

$$L_B = S(R:AC) \quad (8)$$

If a successful cloning was achieved, one expects that $L_A = L_B = 0$, that is, both channels would transmit quantum information perfectly^[1] However, using the chain rule for quantum mutual entropies, it is easy to show that

$$L_A = S(R:BC) = S(R:B) + S(R:C|B) \geq S(R:B) \quad (9)$$

where we have used the property of strong subadditivity of quantum entropies, namely $S(R:C|B) \geq 0$. From Eqs. (4.6), the mutual entropy of the channel $X \rightarrow B$ can be written as $S(R:B) = 2S - L_B$, where $S = S(X)$ is the quantum source entropy. As a consequence, we obtain the *entropic no-cloning inequality*

$$L_A + L_B \geq 2S \quad (10)$$

which implies that the losses L_A and L_B characterizing the two copies cannot vanish simultaneously for $S > 0$ (i.e., when attempting to clone a state that is not fully known). This inequality reflects the impossibility of perfectly copying quantum information and quantifies the balance between the quality of the copies in terms of entropies. It is valid for the cloning of a quantum state in a Hilbert space of arbitrary dimensions. Unfortunately, this entropic no-cloning inequality is not tight. Indeed, by applying the UCM to an arbitrary qubit ($S = 1$), it can be shown that $L_A = L_B = 2 - \log_2(3)/2 = 1.21$ bits, so that $L_A + L_B = 2.42$ bits does not saturate the bound in Eq. (10). In the next Section, we show how to derive a tight inequality for the special case of quantum bits.

3 Pauli Cloning Machines for Quantum Bits

3.1 Characterization of a Pauli Channel Using the Bell States

Consider a quantum bit in an arbitrary state $|\psi\rangle$ which is processed by a Pauli channel. Thus, the qubit is rotated by one of the three Pauli matrices or remains unchanged: it undergoes a phase-flip (σ_z), a bit-flip (σ_x), or their combination ($\sigma_x\sigma_z = -i\sigma_y$) with respective probabilities p_z , p_x , and p_y . (A depolarizing channel corresponds to the special case where $p_x = p_y = p_z$.) It is convenient to describe the operation of such a channel by considering an input maximally entangled with a reference system. Defining the four maximally-entangled states of two qubits (i.e., the Bell states) as

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \quad |\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \quad (11)$$

we note that the *local* action of the Pauli matrices on one of these states, say $|\Phi^+\rangle$, yields the three remaining Bell states, namely^[2]

$$\begin{aligned} (\mathbf{1} \otimes \sigma_z)|\Phi^+\rangle &= |\Phi^-\rangle \\ (\mathbf{1} \otimes \sigma_x)|\Phi^+\rangle &= |\Psi^+\rangle \\ (\mathbf{1} \otimes \sigma_x\sigma_z)|\Phi^+\rangle &= |\Psi^-\rangle \end{aligned} \quad (12)$$

¹ This is a minimal requirement. One could also require the outputs A and B to be independent, for example.

² We use here the convention $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$.

Therefore, if the input qubit X of the Pauli channel is maximally entangled with a reference qubit R , say if their joint state $|\psi\rangle_{RX}$ is the Bell state $|\Phi^+\rangle$, then the joint state of R and the output Y is a mixture of the four Bell states

$$\rho_{RY} = (1-p) |\Phi^+\rangle\langle\Phi^+| + p_z |\Phi^-\rangle\langle\Phi^-| + p_x |\Psi^+\rangle\langle\Psi^+| + p_y |\Psi^-\rangle\langle\Psi^-|, \quad (13)$$

with $p = p_x + p_y + p_z$.

A simple correspondence rule can then be written relating an arbitrary mixture of Bell state and the associated operation on a qubit $|\psi\rangle$ by a Pauli channel. Start from the mixture

$$\rho_{RY} = (1-p) |\Phi^+\rangle\langle\Phi^+| + \sum_{i=1}^3 p_i |\Psi_i\rangle\langle\Psi_i| \quad (14)$$

where $p_1 \leq p_2 \leq p_3$, $p = p_1 + p_2 + p_3$, and $|\Psi_i\rangle$ stand for the three remaining Bell states ranked by increasing weight. It is straightforward to show that the operation on an arbitrary state $|\psi\rangle$ performed by the corresponding channel is

$$\begin{aligned} |\psi\rangle \rightarrow \rho = & (1-p-p_2) |\psi\rangle\langle\psi| + (p_2-p_1) \sigma_1 |\psi_\perp\rangle\langle\psi_\perp| \sigma_1 \\ & + (p_3-p_2) \sigma_3 |\psi\rangle\langle\psi| \sigma_3 + 2(p_1+p_2) \mathbf{1}/2 \end{aligned} \quad (15)$$

where $|\psi_\perp\rangle = -i\sigma_y |\psi^*\rangle = \sigma_x \sigma_z |\psi^*\rangle$ denotes the time-reversed of state $|\psi\rangle$. The four components in the right-hand side of Eq. (15) correspond respectively to the unchanged, (rotated) time-reversed, rotated, and random fraction. It is clear from Eq. (15) that the operation of the channel is *state-independent* only if $p_1 = p_2 = p_3 = p/3$, that is, if the time-reversed and rotated fractions vanish. Then, we have a *depolarizing* channel of probability p , i. e., ρ_{RY} is a Werner state and Eq. (15) becomes

$$|\psi\rangle \rightarrow \rho = (1-4p/3) |\psi\rangle\langle\psi| + (4p/3) \mathbf{1}/2 \quad (16)$$

Thus, the vector characterizing the input qubit in the Bloch sphere is shrunk by a scaling factor $s = 1 - 4p/3$ regardless its orientation, so that the fidelity of the channel, $f = \langle\psi|\rho|\psi\rangle = 1 - 2p/3 = (1+s)/2$, is independent of the input state. Other channels are necessarily *state-dependent*. For example, the “2-Pauli” channel of probability p (i.e., $p_x = p_z = p/2$ and $p_y = 0$) performs the operation

$$\begin{aligned} |\psi\rangle \rightarrow \rho = & (1-3p/2) |\psi\rangle\langle\psi| + (p/2) \sigma_y |\psi_\perp\rangle\langle\psi_\perp| \sigma_y + p \mathbf{1}/2 \\ = & (1-3p/2) |\psi\rangle\langle\psi| + (p/2) |\psi^*\rangle\langle\psi^*| + p \mathbf{1}/2 \end{aligned} \quad (17)$$

while the dephasing channel of probability p (i.e., $p_z = p$ and $p_x = p_y = 0$) simply gives

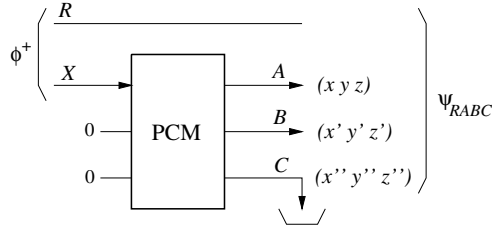
$$|\psi\rangle \rightarrow (1-p) |\psi\rangle\langle\psi| + p \sigma_z |\psi\rangle\langle\psi| \sigma_z \quad (18)$$

3.2 Asymmetric Pauli Cloning Machines

We define an *asymmetric* Pauli cloning machine as a machine whose two outputs, A and B , emerge from distinct Pauli channels [7]. Thus, if the input X of

the cloner is fully entangled with a reference R , i.e., $|\psi\rangle_{RX} = |\Phi^+\rangle$, the density operators ρ_{RA} and ρ_{RB} must then be mixtures of Bell states. Focusing on the first output A , we see that a 4-dimensional Hilbert space is necessary in general to purify ρ_{RA} since we need to accommodate its four (generally nonzero) eigenvalues. The 2-dimensional space of second output qubit B is thus insufficient for this purpose, so that we must introduce an additional system C , which may be viewed as an ancilla or the cloning machine itself. A 2-dimensional space for C is then sufficient, so that we need to consider a single additional qubit C for the cloning machine, as shown in Refs. [3,4]. As a consequence, we are led to consider a 4-qubit system in order to fully describe the PCM, as pictured in Fig. 3. Before cloning, the qubits R and X are in the entangled state $|\Phi^+\rangle$, the two auxiliary qubits being in a prescribed state, e.g., $|0\rangle$. After cloning, the four qubits R , A , B , and C are in a pure state for which ρ_{RA} and ρ_{RB} are mixtures of Bell states (A and B emerge from a Pauli channel). As we shall see, ρ_{RC} happens to be also a mixture of Bell states, so that C can be viewed as a third output emerging from a Pauli channel.

Fig. 3. Pauli cloning machine of input X (initially entangled with a reference R) and outputs A and B . The third output C refers to an ancilla or the cloning machine. The three outputs emerge in general from distinct Pauli channels.



Instead of specifying a PCM by a particular unitary operation acting on the state $|\psi\rangle$ of the input qubit X (together with the two auxiliary qubits in a fixed state $|0\rangle$), it is more convenient to characterize it by the wave function $|\Psi\rangle_{RABC}$ underlying the entanglement of the three outputs with R . So, our goal is to find in general the 4-qubit wave functions that satisfy the requirement that the state of every pair of two qubits is a mixture of the four Bell states. Making use of the Schmidt decomposition of $|\Psi\rangle_{RABC}$ for the bipartite partition RA vs BC , it is clear that this state can be written as a superposition of *double Bell* states

$$|\Psi\rangle_{RA;BC} = \{v |\Phi^+\rangle|\Phi^+\rangle + z |\Phi^-\rangle|\Phi^-\rangle + x |\Psi^+\rangle|\Psi^+\rangle + y |\Psi^-\rangle|\Psi^-\rangle\}_{RA;BC}, \quad (19)$$

where x , y , z , and v are complex amplitudes (with $|x|^2 + |y|^2 + |z|^2 + |v|^2 = 1$). Note that the possible permutations of the Bell states in Eq. (19) are not considered here for simplicity. The above requirement is then satisfied for the qubit pairs RA and BC , that is, $\rho_{RA} = \rho_{BC}$ is of the form of Eq. (13) with $p_x = |x|^2$, $p_y = |y|^2$, $p_z = |z|^2$, and $1 - p = |v|^2$. It is important to note that these double

Bell states for the partition RA vs BC transform into superpositions of double Bell states for the two other possible partitions of the four qubits $RABC$ into two pairs (RB vs AC , RC vs AB). For example, the transformation associated with the partition RB vs AC is

$$\begin{aligned}
|\Phi^+\rangle_{RA} |\Phi^+\rangle_{BC} &= \frac{1}{2} \{ |\Phi^+\rangle|\Phi^+\rangle + |\Phi^-\rangle|\Phi^-\rangle + |\Psi^+\rangle|\Psi^+\rangle + |\Psi^-\rangle|\Psi^-\rangle \}_{RB;AC} \\
|\Phi^-\rangle_{RA} |\Phi^-\rangle_{BC} &= \frac{1}{2} \{ |\Phi^+\rangle|\Phi^+\rangle + |\Phi^-\rangle|\Phi^-\rangle - |\Psi^+\rangle|\Psi^+\rangle - |\Psi^-\rangle|\Psi^-\rangle \}_{RB;AC} \\
|\Psi^+\rangle_{RA} |\Psi^+\rangle_{BC} &= \frac{1}{2} \{ |\Phi^+\rangle|\Phi^+\rangle - |\Phi^-\rangle|\Phi^-\rangle + |\Psi^+\rangle|\Psi^+\rangle - |\Psi^-\rangle|\Psi^-\rangle \}_{RB;AC} \\
|\Psi^-\rangle_{RA} |\Psi^-\rangle_{BC} &= \frac{1}{2} \{ |\Phi^+\rangle|\Phi^+\rangle - |\Phi^-\rangle|\Phi^-\rangle - |\Psi^+\rangle|\Psi^+\rangle + |\Psi^-\rangle|\Psi^-\rangle \}_{RB;AC}
\end{aligned} \tag{20}$$

(For the partition RC vs AB , these expressions are similar up to an overall sign in the transformation of the state $|\Psi^-\rangle_{RA} |\Psi^-\rangle_{BC}$.) This implies that $|\Psi\rangle_{RABC}$ is also a superposition of double Bell states (albeit with different amplitudes) for these two other partitions, which, therefore, also yield mixtures of Bell states when tracing over half of the system. Specifically, for the partition RB vs AC , we obtain

$$|\Psi\rangle_{RB;AC} = \{ v' |\Phi^+\rangle|\Phi^+\rangle + z' |\Phi^-\rangle|\Phi^-\rangle + x' |\Psi^+\rangle|\Psi^+\rangle + y' |\Psi^-\rangle|\Psi^-\rangle \}_{RB;AC} \tag{21}$$

with

$$\begin{aligned}
v' &= (v + z + x + y)/2 \\
z' &= (v + z - x - y)/2 \\
x' &= (v - z + x - y)/2 \\
y' &= (v - z - x + y)/2
\end{aligned} \tag{22}$$

implying that the second output B emerges from a Pauli channel with probabilities $p'_x = |x'|^2$, $p'_y = |y'|^2$, and $p'_z = |z'|^2$. Similarly, the third output C is described by considering the partition RC vs AB ,

$$|\Psi\rangle_{RC;AB} = \{ v'' |\Phi^+\rangle|\Phi^+\rangle + z'' |\Phi^-\rangle|\Phi^-\rangle + x'' |\Psi^+\rangle|\Psi^+\rangle + y'' |\Psi^-\rangle|\Psi^-\rangle \}_{RC;AB} \tag{23}$$

with

$$\begin{aligned}
v'' &= (v + z + x - y)/2 \\
z'' &= (v + z - x + y)/2 \\
x'' &= (v - z + x + y)/2 \\
y'' &= (v - z - x - y)/2
\end{aligned} \tag{24}$$

Thus, Eqs. (22) and (24) relate the amplitudes of the double Bell states for the three possible partitions of the four qubits into two pairs, and thereby specify the entire set of asymmetric Pauli cloning machines.

3.3 No-cloning Uncertainty Relation for Quantum Bits

The complementarity between the two copies produced by a PCM can be shown to result in general from an *uncertainty principle*, much like that associated with Fourier transforms. In order to show this, let us construct the two-dimensional discrete function $a_{m,n}$ with $m, n = 0, 1$: let $a_{0,0} = v$, $a_{0,1} = z$, $a_{1,0} = x$, and $a_{1,1} = y$. Thus, output A emerges from a Pauli channel characterized by the probability distribution $|a_{m,n}|^2$, where $p_z = |a_{0,1}|^2$, $p_x = |a_{1,0}|^2$, $p_y = |a_{1,1}|^2$, and $|a_{0,0}|^2$ is simply the probability that the qubit remains unchanged. Similarly, output B can be characterized by a two-dimensional function $b_{m,n}$ defined as $b_{0,0} = v'$, $b_{0,1} = z'$, $b_{1,0} = x'$, and $b_{1,1} = y'$. Using this notation, it appears that Eq. (22) is simply a two-dimensional discrete Fourier transform,

$$b_{m,n} = \frac{1}{2} \sum_{x=0}^1 \sum_{y=0}^1 (-1)^{nx+my} a_{x,y} \quad (25)$$

This emphasizes that if output A is close to perfect ($a_{m,n}$ is a peaked function) then output B is very noisy ($b_{m,n}$ is a flat function), and conversely. Consequently, the probability distributions $|a_{m,n}|^2$ and $|b_{m,n}|^2$ characterizing the channels underlying the two outputs A and B cannot have a variance simultaneously tending to zero, giving rise to an uncertainty principle which governs the tradeoff between the quality of the copies.

To illustrate this no-cloning uncertainty principle, let us consider the class of asymmetric PCMs whose outputs A and B emerge from (distinct) *depolarizing* channels. Assume that the first output A emerges from a depolarizing channel of probability $p = 3|x|^2$, i.e.,

$$\rho_{RA} = |v|^2 |\Phi^+\rangle\langle\Phi^+| + |x|^2 (|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|) , \quad (26)$$

with $|v|^2 + 3|x|^2 = 1$. Then, from Eq. (22), we have $v' = (v + 3x)/2$ and $x' = (v - x)/2$, resulting in

$$\rho_{RB} = \frac{|v + 3x|^2}{4} |\Phi^+\rangle\langle\Phi^+| + \frac{|v - x|^2}{4} (|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|) . \quad (27)$$

Thus, the second output B also emerges from a depolarizing channel of probability $p' = 3|x'|^2 = \frac{3}{4}|v - x|^2$, implying that both outputs of this asymmetric PCM are state-independent and simply correspond to a different scaling of the vector characterizing the input qubit in the Bloch sphere. (The third output C emerges in general from a different Pauli channel.) The relation between the parameters x and x' characterizing the two outputs can be written as

$$|x|^2 + \text{Re}(x^* x') + |x'|^2 = \frac{1}{4} \quad (28)$$

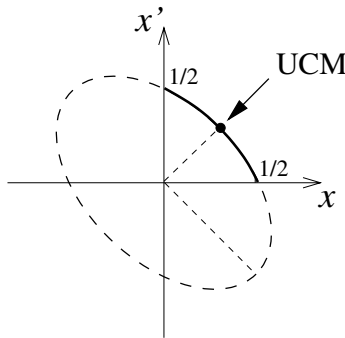
Clearly, the best cloning (minimum values for $|x|$ and $|x'|$) is achieved when the cross term is the largest in magnitude, that is, when x and x' have the same phase. For simplicity, we assume here that x and x' are real and positive.

Consequently, the tradeoff between the quality of the two copies can be described by the *no-cloning uncertainty relation*

$$x^2 + xx' + x'^2 \geq \frac{1}{4}, \quad (29)$$

where the copying error is measured by the probability of the depolarizing channel underlying each output, i.e., $p = 3x^2$ and $p' = 3x'^2$ (with $x, x' \geq 0$). Equation (29) corresponds to the domain in the (x, x') -space located outside an ellipse whose semiminor axis, oriented in the direction $(1, 1)$, is $1/\sqrt{6}$, as shown in Fig. 4. (The semimajor axis is $1/\sqrt{2}$.) The origin in this space corresponds to a (nonexisting) cloner whose two outputs would be perfect $p = p' = 0$, while to distance to origin measures $(p+p')/3$. The ellipse characterizes the ensemble of values for p and p' that can actually be achieved with a PCM. It intercepts its minor axis at $(1/\sqrt{12}, 1/\sqrt{12})$, which corresponds to the universal cloning machine (UCM), i.e., $p = p' = 1/4$, as discussed below. This point is the closest to the origin (i.e., the cloner with minimum $p+p'$), and characterizes in this sense the best possible copying. The UCM is the only symmetric cloner belonging to the class of PCM considered here (i.e., cloners whose outputs are depolarizing channels); other symmetric cloners will be considered in Sec. 3.4. The ellipse crosses the x -axis at $(1/2, 0)$, which describes the situation where the first output emerges from a 100%-depolarizing channel ($p = 3/4$) while the second emerges from a perfect channel ($p' = 0$). Of course, $(0, 1/2)$ corresponds to the symmetric situation.

Fig. 4. Ellipse delimiting the best quality of the two outputs of an asymmetric PCM that can be achieved simultaneously (only the quadrant $x, x' \geq 0$ is of interest here). The outputs emerge from depolarizing channels of probability $p = 3x^2$ and $p' = 3x'^2$. Any close-to-perfect cloning characterized by a point inside the ellipse is forbidden.



The dimensional argument used in Sec. 3.2 strongly suggests that the imperfect cloning achieved by an asymmetric PCM as described above is optimal: a single additional qubit C for the cloner is sufficient to perform the best cloning, i.e., to achieve the minimum p and p' for a fixed ratio p/p' . (This is proven rigorously for the special case $p = p'$ in Ref. [4]). Also, introducing a phase difference

between x and x' results in a set of PCMs characterized by an ellipse that is less eccentric and tends to a circle of radius $1/2$ for a phase difference of $\pi/2$. Consequently, the no-cloning inequality (29) is saturated when x and x' have the same (or opposite) phase. The domain inside the ellipse corresponds then to the values for p and p' that cannot be achieved simultaneously, reflecting the impossibility of close-to-perfect cloning, and Eq. (29) is the tightest no-cloning bound that can be written for a qubit.

The no-cloning uncertainty relation for qubits can also be reexpressed as

$$a^2 + ab + b^2 \geq 1, \quad (30)$$

where $a^2 = 4p/3 = 4x^2$ and $b^2 = 4p'/3 = 4x'^2$ denote the *depolarizing fraction* of the channels leading to outputs A and B , respectively.³ The corresponding ellipse crosses the axes at $(0, 1)$ and $(1, 0)$, while the closest point to the origin is $(\sqrt{1/3}, \sqrt{1/3})$. Indeed, the outputs of the UCM emerge from two channels whose depolarizing fraction is $1/3$. The uncertainty relation associated with the cloning of N -dimensional quantum states (instead of qubits) is investigated in Ref. [11]. It is shown there that the cross-term in Eq. (30) becomes $2ab/N$, implying that the ellipse tends to a circle of radius one at the limit of a large dimension N .

3.4 Symmetric Pauli Cloning Machines

Consider now the class of symmetric PCMs that have both outputs emerging from a *same* Pauli channel, i.e., $\rho_{RA} = \rho_{RB}$. Using Eq. (22), we obtain the conditions

$$\begin{aligned} |v|^2 &= |v + z + x + y|^2/4 \\ |z|^2 &= |v + z - x - y|^2/4 \\ |x|^2 &= |v - z + x - y|^2/4 \\ |y|^2 &= |v - z - x + y|^2/4 \end{aligned} \quad (31)$$

which yields

$$v = x + y + z, \quad (32)$$

where x , y , z , and v are assumed to be real. Equation (32), together with the normalization condition, describes a two-dimensional surface in a space where each point (x, y, z) ⁴ represents a Pauli channel of parameters $p_x = x^2$, $p_y = y^2$, and $p_z = z^2$. This surface,

$$x^2 + y^2 + z^2 + xy + xz + yz = \frac{1}{2}, \quad (33)$$

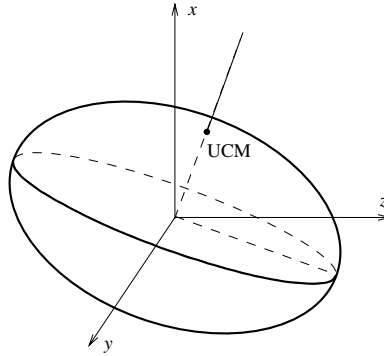
is an oblate ellipsoid E with symmetry axis along the direction $(1, 1, 1)$, as shown in Fig. 5. The semiminor axis (or polar radius) is $1/2$ while the semimajor axis

³ The input qubit is replaced by a random qubit with probability a^2 (b^2) or left unchanged with probability $1 - a^2$ ($1 - b^2$) in the channel leading to output A (B).

⁴ We only consider here the first octant $x, y, z \geq 0$.

(or equatorial radius) is 1. In this representation, the distance to the origin is $p_x + p_y + p_z$, so that the pole $(1/\sqrt{12}, 1/\sqrt{12}, 1/\sqrt{12})$ of this ellipsoid—the closest point to the origin—corresponds to the special case of a depolarizing channel of probability $p = 1/4$. Thus, this particular PCM reduces to the UCM. This simply illustrates that the requirement of having an optimal cloning (minimum $p_x + p_y + p_z$) implies that the cloner is state-independent ($p_x = p_y = p_z$).

Fig. 5. Oblate ellipsoid representing the class of symmetric PCMs whose two outputs emerge from the same Pauli channel of parameters $p_x = x^2$, $p_y = y^2$, and $p_z = z^2$ (only the octant $x, y, z \geq 0$ is considered here). The pole of this ellipsoid corresponds to the UCM. The capacity of a Pauli channel that lies outside this ellipsoid must be vanishing.



The parametric equations of ellipsoid E can be written as

$$\begin{aligned} x &= \sqrt{\frac{1}{12}} \cos(\theta) + \sqrt{\frac{2}{3}} \sin(\theta) \cos(\phi) \\ y &= \sqrt{\frac{1}{12}} \cos(\theta) + \sqrt{\frac{2}{3}} \sin(\theta) \cos(\phi + 2\pi/3) \\ z &= \sqrt{\frac{1}{12}} \cos(\theta) + \sqrt{\frac{2}{3}} \sin(\theta) \cos(\phi + 4\pi/3) \end{aligned} \quad (34)$$

where the polar angle θ measures the “distance” from a depolarizing channel ($\theta = 0$ implies $p_x = p_y = p_z$), while the azimuthal angle ϕ characterizes the distribution among p_x , p_y , and p_z .

3.5 Universal Cloning Machine

The optimal symmetric PCM (i.e., the UCM) can be obtained alternatively by requiring that the two outputs A and B of a symmetric cloner are maximally independent. Using Eqs. (24) and (32), we obtain $v'' = x + z$, $z'' = y + z$, $x'' = x + y$, and $y'' = 0$. Therefore, we have

$$\rho_{RC} = \rho_{AB} = |x + z|^2 |\Phi^+\rangle\langle\Phi^+| + |y + z|^2 |\Phi^-\rangle\langle\Phi^-| + |x + y|^2 |\Psi^+\rangle\langle\Psi^+|. \quad (35)$$

showing that the third output C emerges from a Pauli channel with vanishing p_y . Thus, we want to maximize the joint von Neumann entropy of the two outputs A and B ,

$$S(AB) = -\text{Tr}(\rho_{AB} \log \rho_{AB}) = H [|x+z|^2, |y+z|^2, |x+y|^2] \quad (36)$$

with $H[\cdot]$ denoting the Shannon entropy. It is easy to see that the solution with $x, y, z \geq 0$ that maximizes $S(AB)$ is $x = y = z$, that is, the Pauli channel underlying outputs A and B reduces to a depolarizing channel. Using Eq. (33), we get $x = y = z = 1/\sqrt{12}$, so that the wave function underlying the UCM is

$$\begin{aligned} |\Psi\rangle_{RA;BC} &= \sqrt{\frac{3}{4}} |\Phi^+\rangle_{RA} |\Phi^+\rangle_{BC} \\ &+ \sqrt{\frac{1}{12}} \{ |\Phi^-\rangle |\Phi^-\rangle + |\Psi^+\rangle |\Psi^+\rangle + |\Psi^-\rangle |\Psi^-\rangle \}_{RA;BC} \end{aligned} \quad (37)$$

Consequently

$$\rho_{RA} = \rho_{RB} = \frac{3}{4} |\Phi^+\rangle \langle \Phi^+| + \frac{1}{12} (|\Phi^-\rangle \langle \Phi^-| + |\Psi^+\rangle \langle \Psi^+| + |\Psi^-\rangle \langle \Psi^-|) \quad (38)$$

confirming that A and B emerge both from a depolarizing channel with $p = 1/4$,

$$|\psi\rangle \rightarrow \frac{2}{3} |\psi\rangle \langle \psi| + \frac{1}{3} (\mathbf{1}/2) \quad (39)$$

so that the scaling factor is $s = 1 - 4p/3 = 2/3$ while the fidelity of cloning is $f = 1 - 2p/3 = 5/6$ [3]. For the partition RC vs AB , we obtain

$$|\Psi\rangle_{RC;AB} = \sqrt{\frac{1}{3}} \{ |\Phi^+\rangle |\Phi^+\rangle + |\Phi^-\rangle |\Phi^-\rangle + |\Psi^+\rangle |\Psi^+\rangle \}_{RC;AB} \quad (40)$$

implying that the 4-qubit wave function is symmetric under the interchange of A and B (or R and C). It is easy to check that the unitary transformation which implements the UCM [3] is

$$\begin{aligned} |0\rangle_X |00\rangle &\rightarrow \sqrt{\frac{2}{3}} |00\rangle_{AB} |0\rangle_C + \sqrt{\frac{1}{3}} |\Psi^+\rangle_{AB} |1\rangle_C \\ |1\rangle_X |00\rangle &\rightarrow \sqrt{\frac{2}{3}} |11\rangle_{AB} |1\rangle_C + \sqrt{\frac{1}{3}} |\Psi^+\rangle_{AB} |0\rangle_C \end{aligned} \quad (41)$$

Indeed, using Eq. (41), we have

$$\begin{aligned} |\Phi^+\rangle_{RX} |00\rangle &\rightarrow \sqrt{\frac{1}{3}} (|00\rangle_{AB} |00\rangle_{RC} + |11\rangle_{AB} |11\rangle_{RC}) \\ &+ \sqrt{\frac{1}{6}} |\Psi^+\rangle_{AB} (|01\rangle_{RC} + |10\rangle_{RC}) \end{aligned} \quad (42)$$

so that the initial state of X (maximally entangled with the reference R) is transformed into the wave function Eq. (40). The latter implies

$$\rho_{RC} = \rho_{AB} = \frac{1}{3} (|\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+|) \quad (43)$$

showing that the joint entropy of the two outputs is maximum (remember that the singlet $|\Psi^-\rangle$ component must vanish). Thus, the third output of the UCM emerges from a 2-Pauli channel of probability $2/3$. Using Eq. (17), we see that the corresponding operation on an arbitrary state $|\psi\rangle$ is

$$|\psi\rangle \rightarrow \frac{1}{3} |\psi^*\rangle\langle\psi^*| + \frac{2}{3} (1/2) \quad (44)$$

as noted in Ref. [12].

3.6 Related Bound on the Capacity of the Pauli Channel

An interesting application of the UCM is that it can be used to establish an upper bound on the quantum capacity C of the depolarizing channel, namely $C = 0$ at $p = 1/4$ [4]. Extending this result, we show here that the class of symmetric PCMs characterized by Eq. (33) puts a limit on the quantum capacity of Pauli channels. Consider a PCM whose outputs emerge from a Pauli channel of probabilities p_x, p_y , and p_z . Applying an error-correcting scheme separately on each output of the cloning machine (obviously of the other output) would lead to a violation of the no-cloning theorem if the capacity $C(p_x, p_y, p_z)$ was nonzero. Since C is a nonincreasing function of p_x, p_y , and p_z , for $p_x, p_y, p_z \leq 1/2$ (i.e., adding noise to a channel cannot increase its capacity), we conclude that

$$C(p_x, p_y, p_z) = 0 \quad \text{if } (x, y, z) \notin E \quad (45)$$

that is, the quantum capacity is vanishing for any Pauli channel that lies *outside* the ellipsoid E . In particular, Eq. (33) implies that the quantum capacity vanishes for (i) a depolarizing channel with $p = 1/4$ ($p_x = p_y = p_z = 1/12$) [4]; (ii) a “2-Pauli” channel with $p = 1/3$ ($p_x = p_z = 1/6, p_y = 0$); and (iii) a dephasing channel with $p = 1/2$ ($p_x = p_y = 0, p_z = 1/2$). Furthermore, using the fact that C cannot be superadditive for a convex combination of a perfect and a noisy channel [13], an upper bound on C can be written using a linear interpolation between the perfect channel $(0, 0, 0)$ and any Pauli channel lying on E :

$$C \leq 1 - 2(x^2 + y^2 + z^2 + xy + xz + yz) . \quad (46)$$

Note that another class of symmetric PCMs can be found by requiring $\rho_{RA} = \rho_{RC}$, i.e., considering C as the second output and B as the cloning machine. This requirement implies $v = x - y + z$ rather than Eq. (32), which gives rise to the reflection of E with respect to the xz -plane, i.e. $y \rightarrow -y$. It does not change the above bound on C because this class of PCMs has noisier outputs in the first octant $x, y, z \geq 0$.

3.7 Quantum Triplicators Based on the PCM

Let us turn to the fully symmetric PCMs that have *three* outputs emerging from the *same* Pauli channel, i.e., $\rho_{RA} = \rho_{RB} = \rho_{RC}$, which corresponds to a family of (non-optimal) quantum *triplicating* machines. The requirement $\rho_{RA} = \rho_{RC}$ implies $v = x - y + z$, which, together with Eq. (32), yields the conditions $(v = x + z) \wedge (y = 0)$. Incidentally, we notice that if *all* pairs are required to be in the *same* mixture of Bell states, this mixture cannot have a singlet $|\Psi^-\rangle$ component. The outputs of the corresponding triplicators emerge therefore from a “2-Pauli” channel ($p_y = 0$), so that these triplicators are *state-dependent*, in contrast with the one considered in Ref. [5]. (For describing a *state-independent* triplicator, a 6-qubit wave function should be used, that is, the cloner should consist of 2 qubits.) These triplicators are represented by the intersection of E with the xz -plane, that is, the ellipse

$$x^2 + z^2 + xz = \frac{1}{2}, \quad (47)$$

whose semiminor axis is $1/\sqrt{3}$ [oriented along the direction $(1, 1)$] and semimajor axis is 1. The intersection of this ellipse with its semiminor axis ($x = z = 1/\sqrt{6}$) corresponds to the 4-qubit wave function

$$|\Psi\rangle_{abcd} = \frac{2}{\sqrt{6}}|\Phi^+\rangle|\Phi^+\rangle + \frac{1}{\sqrt{6}}|\Phi^-\rangle|\Phi^-\rangle + \frac{1}{\sqrt{6}}|\Psi^+\rangle|\Psi^+\rangle, \quad (48)$$

which is symmetric under the interchange of any two qubits and maximizes the 2-bit entropy (or minimizes the mutual entropy between any two outputs of the triplicator, making them maximally independent). Equation (48) thus characterizes the best triplicator of this ensemble, whose three outputs emerge from a “2-Pauli” channel with $p = 1/3$ ($p_x = p_z = 1/6$). According to Eq. (17), the (state-dependent) operation of this triplicator on an arbitrary qubit can be written as

$$|\psi\rangle \rightarrow \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{6}|\psi^*\rangle\langle\psi^*| + \frac{1}{3}(\mathbf{1}/2). \quad (49)$$

If $|\psi\rangle$ is real, Eq. (49) reduces to the triplicator that was considered in Ref. [12]. The fidelity of cloning is then the same as for the UCM, i.e., $f = 5/6$, regardless the input state (provided it is real).

4 Conclusion

We have derived an entropic no-cloning inequality, characterizing the balance between the losses of the quantum channels leading to the two outputs of a cloning machine. This inequality is valid for states in a Hilbert space of arbitrary dimension and implies that the losses of the two channels cannot vanish simultaneously as long as the source entropy is non-zero. Then, we have focused on the cloning of quantum bits and defined a family of Pauli cloning machines (PCM). These cloners, whose outputs emerge from two distinct Pauli channels, generalize the

universal cloning machine (UCM). The UCM is a special case (symmetric and state-independent) of a PCM. The asymmetric PCMs allowed us to derive a tight no-cloning uncertainty relation for quantum bits, characterizing the impossibility of (perfectly) copying due to quantum mechanics. Furthermore, we have shown that the tradeoff between the quality of the two outputs of a PCM results in general from an uncertainty principle akin to the complementarity between position and momentum. Specifically, the probability distributions characterizing the channels associated with the two outputs cannot be peaked simultaneously, this complementarity being simply governed by the relationship between a function and its Fourier transform. Finally, using a class of symmetric PCMs, we have established an upper bound on the quantum capacity of the Pauli channel.

We acknowledge C. Adami for many useful discussions. This work was supported in part by the NSF under Grant Nos. PHY 94-12818 and PHY 94-20470, and by a grant from DARPA/ARO through the QUIC Program (#DAAH04-96-1-3086). N.J.C. is Collaborateur scientifique of the Belgian National Fund for Scientific Research.

References

1. D. Dieks, Phys. Lett. **92A**, 271 (1982).
2. W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
3. V. Buzek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).
4. D. Bruss, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998).
5. N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
6. N. Gisin and B. Huttner, Phys. Lett. A **228**, 13 (1997).
7. N. J. Cerf, Los Alamos e-print quant-ph/9803058.
8. C. Adami and N. J. Cerf, Phys. Rev. A **56**, 3470 (1997).
9. N. J. Cerf, Phys. Rev. A **57**, 3330 (1998).
10. B. Schumacher, Phys. Rev. A **54**, 2614 (1996); B. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996).
11. N. J. Cerf, Los Alamos e-print quant-ph/9805024.
12. V. Buzek, S. L. Braunstein, M. Hillery, and D. Bruss, Phys. Rev. A **56**, 3446 (1997).
13. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

Universal Optimal Cloning of Qubits and Quantum Registers

V. Bužek¹ and M. Hillery²

¹ Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 842 28
Bratislava, Slovakia

² Department of Physics and Astronomy, Hunter College, CUNY, 695 Park Avenue,
New York, NY 10021, USA

Abstract. We review our recent work on the universal (i.e. input state independent) optimal quantum copying (cloning) of qubits. We present unitary transformations which describe the optimal cloning of a qubit, and we present the corresponding quantum logic network. We also present a network for an optimal quantum copying “machine” (transformation) which produces $N + 1$ identical copies from the original qubit. Here again the quality (fidelity) of the copies does not depend on the state of the original and is only a function of the number of copies, N . In addition, we present the machine which universally clones states of quantum objects in arbitrary-dimensional Hilbert spaces. In particular, we discuss universal cloning of quantum registers.

Keywords: quantum cloning, quantum logic networks, inseparability

1 Introduction

The most fundamental difference between classical and quantum information is that while classical information can be copied perfectly, quantum information cannot. In particular, it follows from the *no-cloning theorem* [1] (see also [23]) that one cannot create a perfect duplicate of an *arbitrary* qubit while preserving the state of the original qubit. In the quantum *teleportation* (see [4] and references therein), one can create a perfect copy of the original qubit but this will be at the expense of the *complete* destruction of information encoded in the original qubit. In contrast, the main goal of quantum copying is to produce a copy of the original qubit which is as close as possible to the original state while the output state of the original qubit is minimally disturbed.

We have shown recently [5,6] that if one is only interested in producing *imperfect* copies, then it is possible to make quantum clones of the original qubit. For example, the copy machine considered by Wootters and Zurek [1] in their proof of the no-cloning theorem produces two identical copies at its output, but the quality of these copies depends upon the input state. They are perfect for the basis vectors, which we denote as $|0\rangle$ and $|1\rangle$, but, because the copying process destroys the off-diagonal information of the input density matrix, they are poor for input states of the form $(|1\rangle + e^{i\varphi}|0\rangle)/\sqrt{2}$, where φ is arbitrary. We have introduced [5,6,7] a different copying machine, the Universal Quantum

Copying Machine (UQCM), which produces two identical copies whose quality is *independent* of the input state. In addition, its performance is, on average, better than that of the Wootters-Zurek machine, and the action of the machine simply scales the expectation values of relevant observables. This UQCM was shown to be optimal, in the sense that it maximizes the average fidelity between the input and output qubits, by Gisin and Massar [8] and by Bruß et al. [9]. Gisin and Massar have also been able to find copying transformations which produce N copies from M originals (where $N > M$) [8]. In addition, we have proposed quantum logic networks for quantum copying (cloning) machines [7,10], and bounds have been placed on how good copies can be [11,12].

In this paper we will firstly review our original ideas on the universal quantum copying of a single qubit (Section II). In Section III we will present a quantum network describing the UQCM. Secondly, in Section IV we will introduce the copying machine which produces $N + 1$ identical copies from the original qubit and present a quantum network which implements it. The quality (fidelity) of the copies does not depend on the state of the original and is only a function of a number N of copies produced. This machine is formally described by the same unitary transformation recently introduced by Gisin and Massar [8]. In Section V we will analyze the properties of multiply cloned qubits. Thirdly, in Section VI we show how quantum registers (i.e. systems composed of many entangled qubits) can be universally cloned. One approach is to use single-qubit copiers to individually (locally) copy each qubit. We have shown earlier [13] that in the case of two qubits this local copying will preserve some of the quantum correlations between qubits, but as we will show, it does not make a particularly good copy of the two-qubit state. As an alternative we propose a copy machine which universally clones quantum states in arbitrary-dimensional Hilbert spaces. This allow us to discuss the cloning of quantum registers.

2 Universal Quantum Copying Machine

Let us assume we want to copy an arbitrary pure state $|\Psi\rangle_{a_0}$ which in a particular basis $\{|0\rangle_{a_0}, |1\rangle_{a_0}\}$ is described by the state vector $|\Psi\rangle_{a_0}$

$$|\Psi\rangle_{a_0} = \alpha|0\rangle_{a_0} + \beta|1\rangle_{a_0}; \quad \alpha = \sin \vartheta/2e^{i\varphi}; \quad \beta = \cos \vartheta/2. \quad (1)$$

The two numbers which characterize the state (1) can be associated with the “amplitude” $|\alpha|$ and the “phase” φ of the qubit. Even though ideal copying, i.e., the transformation $|\Psi\rangle_{a_0} \longrightarrow |\Psi\rangle_{a_0}|\Psi\rangle_{a_1}$ is prohibited by the laws of quantum mechanics for an *arbitrary* state (1), it is still possible to design quantum copiers which operate reasonably well. In particular, the UQCM [5] is specified by the following conditions:

(i) The state of the original system and its quantum copy at the output of the quantum copier, described by density operators $\hat{\rho}_{a_0}^{(out)}$ and $\hat{\rho}_{a_1}^{(out)}$, respectively, are identical, i.e.,

$$\hat{\rho}_{a_0}^{(out)} = \hat{\rho}_{a_1}^{(out)} \quad (2)$$

(ii) If no *a priori* information about the *in*-state of the original system is available, then it is reasonable to require that *all* pure states should be copied equally well. One way to implement this assumption is to design a quantum copier such that the distances between density operators of each system at the output ($\hat{\rho}_{a_j}^{(out)}$ where $j = 0, 1$) and the ideal density operator $\hat{\rho}^{(id)}$ which describes the *in*-state of the original mode are input-state independent. Quantitatively this means that if we employ the Bures distance [14]

$$d_B(\hat{\rho}_1, \hat{\rho}_2) = \sqrt{2} \left(1 - \text{Tr} \sqrt{\hat{\rho}_1^{1/2} \hat{\rho}_2 \hat{\rho}_1^{1/2}} \right)^{1/2}, \quad (3)$$

as a measure of distance between two operators, then the quantum copier should be such that

$$d_B(\hat{\rho}_{a_j}^{(out)}; \hat{\rho}_{a_j}^{(id)}) = \text{const.}; \quad j = 0, 1. \quad (4)$$

(iii) Finally, we would also like to require that the copies are as close as possible to the ideal output state, which is, of course, just the input state. This means that we want our quantum copying transformation to minimize the distance between the output state $\hat{\rho}_{a_j}^{(out)}$ of the copied qubit and the ideal state $\hat{\rho}_{a_j}^{(id)}$. The distance is minimized with respect to all possible unitary transformations U acting on the Hilbert space \mathcal{H} of two qubits and the quantum copying machine (i.e., $\mathcal{H} = \mathcal{H}_{a_0} \otimes \mathcal{H}_{a_1} \otimes \mathcal{H}_x$)

$$d_B(\hat{\rho}_{a_j}^{(out)}; \hat{\rho}_{a_j}^{(id)}) = \min \left\{ d_B^{(U)}(\hat{\rho}_{a_j}^{(out)}; \hat{\rho}_{a_j}^{(id)}); \forall U \right\}; \quad (j = 0, 1). \quad (5)$$

Originally, the UQCM was found by analyzing a transformation which contained two free parameters, and then determining them by demanding that condition (ii) be satisfied, and that the distance between the two-qubit output density matrix and the ideal two-qubit output be input-state independent. That the UQCM machine obeys the condition (5) has only been shown recently [8,9].

The unitary transformation which implements the UQCM [5] is given by

$$\begin{aligned} |0\rangle_{a_0} |Q\rangle_x &\rightarrow \sqrt{\frac{2}{3}} |00\rangle_{a_0 a_1} |\uparrow\rangle_x + \sqrt{\frac{1}{3}} |+\rangle_{a_0 a_1} |\downarrow\rangle_x \\ |1\rangle_{a_0} |Q\rangle_x &\rightarrow \sqrt{\frac{2}{3}} |11\rangle_{a_0 a_1} |\downarrow\rangle_x + \sqrt{\frac{1}{3}} |+\rangle_{a_0 a_1} |\uparrow\rangle_x, \end{aligned} \quad (6)$$

where $|+\rangle_{a_0 a_1} = (|10\rangle_{a_0 a_1} + |01\rangle_{a_0 a_1})/\sqrt{2}$. This transformation satisfies the conditions (2,5). The system labelled by a_0 is the original (input) qubit, while the other system a_1 represents the qubit onto which the information is copied. This qubit is prepared initially in the state $|0\rangle_{a_1}$ (it plays the role that a blank piece of paper plays in a copier). States of the copy machine are labelled by x . The state space of the copy machine is two dimensional, and we assume that it is always in the same state $|Q\rangle_x$ initially. If the original qubit is in the superposition state

(II) then the reduced density operators of both copies at the output are equal [see condition (2)] and they can be expressed as

$$\hat{\rho}_{a_j}^{(out)} = \frac{5}{6} |\Psi\rangle_{a_j} \langle \Psi| + \frac{1}{6} |\Psi_\perp\rangle_{a_j} \langle \Psi_\perp|, \quad j = 0, 1 \quad (7)$$

where $|\Psi_\perp\rangle_{a_j} = \beta^* |0\rangle_{a_j} - \alpha^* |1\rangle_{a_j}$, is the state orthogonal to $|\Psi\rangle_{a_j}$. This implies that the copy contains 5/6 of the state we want and 1/6 of the one we do not.

The density operator $\rho_{a_j}^{(out)}$ given by Eq. (7) can be rewritten in a “scaled” form:

$$\hat{\rho}_{a_j}^{(out)} = s_j \hat{\rho}_{a_j}^{(id)} + \frac{1 - s_j}{2} \hat{1}; \quad j = 0, 1, \quad (8)$$

which guarantees that the distance (3) is input-state independent, i.e. the condition (4) is automatically fulfilled. The scaling factor in Eq. (8) is $s_j = 2/3$ ($j = 0, 1$).

The Bures distance $d_B(\hat{\rho}_{a_j}^{(out)}; \hat{\rho}_{a_j}^{(id)})$ ($j = 0, 1$) between the output qubit and the ideal qubit is constant and it reads

$$d_B(\hat{\rho}_{a_j}^{(out)}; \hat{\rho}_{a_j}^{(id)}) = \sqrt{2} \left(1 - \sqrt{\frac{5}{6}} \right). \quad (9)$$

We note, that the idle qubit after the copying is performed is in a state

$$\hat{\rho}_{b_1}^{(out)} = \frac{1}{3} \left(\hat{\rho}_{b_1}^{(id)} \right)^T + \frac{1}{3} \hat{1}, \quad (10)$$

where the superscript T denotes the transpose.

We stress once again that the UQCM copies all input states with the same quality and therefore is suitable for copying when no *a priori* information about the state of the original qubit is available. This corresponds to a uniform prior probability distribution on the state space of a qubit (Poincare sphere). Correspondingly, one can measure the quality of copies by the fidelity \mathcal{F} , which is equal to the mean overlap between a copy and the input state [8]

$$\mathcal{F} = \int d\Omega_{a_j} \langle \Psi | \hat{\rho}_{a_j}^{(out)} | \Psi \rangle_{a_j}, \quad (11)$$

where $\int d\Omega = \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta \sin \vartheta / 4\pi$. It is easy to show that the relation between the fidelity \mathcal{F} and the scaling factor is $s = 2\mathcal{F} - 1$.

3 Copying Network

In what follows we show how, with simple quantum logic gates, we can copy quantum information encoded in the original qubit onto other qubits. The copying procedure can be understood as a “spread” of information via a “controlled” entanglement between the original qubit and the copy qubits. This controlled

entanglement is implemented by a sequence of controlled-NOT operations operating on the original qubit and the copy qubits which are initially prepared in a specific state.

In designing a network for the UQCM we first note that since the state space of the copy machine itself is two dimensional, we can consider it to be an additional qubit. Our network, then, will take 3 input qubits (one for the input, one which becomes a copy, and one for the machine) and transform them into 3 output qubits. In what follows we will denote the quantum copier qubit as b_1 rather than x . The operation of this network is such, that in order to transfer information from the original a_0 qubit to the target qubit a_1 we will need one *idle* qubit b_1 which plays the role of quantum copier.

Before proceeding with the network itself let us specify the one and two-qubit gates from which it will be constructed. Firstly, we define a single-qubit rotation $\hat{R}_j(\theta)$ which acts on the basis vectors of qubits as

$$\hat{R}_j(\theta)|0\rangle_j = \cos \theta|0\rangle_j + \sin \theta|1\rangle_j; \quad \hat{R}_j(\theta)|1\rangle_j = -\sin \theta|0\rangle_j + \cos \theta|1\rangle_j. \quad (12)$$

We also will utilize a two-qubit operator (a two-bit quantum gate), the so-called controlled-NOT gate, which has as its inputs a control qubit and a target qubit. The operator which implements this gate, \hat{P}_{kl} , acts on the basis vectors of the two qubits as follows (k denotes the control qubit and l the target):

$$\begin{aligned} \hat{P}_{kl}|0\rangle_k|0\rangle_l &= |0\rangle_k|0\rangle_l; & \hat{P}_{kl}|0\rangle_k|1\rangle_l &= |0\rangle_k|1\rangle_l; \\ \hat{P}_{kl}|1\rangle_k|0\rangle_l &= |1\rangle_k|1\rangle_l; & \hat{P}_{kl}|1\rangle_k|1\rangle_l &= |1\rangle_k|0\rangle_l. \end{aligned} \quad (13)$$

We can decompose the quantum copier network into two parts. In the first part the copy (a_1) and the idle (b_1) qubits are prepared in a specific state $|\Psi\rangle_{a_1 b_1}^{(prep)}$. Then in the second part, the information from the original qubit a_0 is *redistributed* among the three qubits. That is, the action of the quantum copier can be described as a sequence of two unitary transformations

$$|\Psi\rangle_{a_0}^{(in)}|0\rangle_{a_1}|0\rangle_{b_1} \longrightarrow |\Psi\rangle_{a_0}^{(in)}|\Psi\rangle_{a_1 b_1}^{(prep)} \longrightarrow |\Psi\rangle_{a_0 a_1 b_1}^{(out)}. \quad (14)$$

3.1 Preparation of Quantum Copier

Let us first look at the preparation stage. Prior to any interaction with the input qubit we have to prepare the two quantum copier qubits (a_1 and b_1) in a specific state $|\Psi\rangle_{a_1 b_1}^{(prep)}$

$$|\Psi\rangle_{a_1 b_1}^{(prep)} = \frac{1}{\sqrt{6}} (2|00\rangle_{a_1 b_1} + |01\rangle_{a_1 b_1} + |11\rangle_{a_1 b_1}), \quad (15)$$

which can be prepared by a simple quantum network with two controlled-NOTs \hat{P}_{kl} and three rotations $\hat{R}(\theta_j)$, i.e.

$$|\Psi\rangle_{a_1 b_1}^{(prep)} = \hat{R}_{a_1}(\theta_3)\hat{P}_{b_1 a_1}\hat{R}_{b_1}(\theta_2)\hat{P}_{a_1 b_1}\hat{R}_{a_1}(\theta_1)|0\rangle_{a_1}|0\rangle_{b_1}, \quad (16)$$

with the rotation angles defined as [7]

$$\cos 2\theta_1 = \frac{1}{\sqrt{5}}; \quad \cos 2\theta_2 = \frac{\sqrt{5}}{3}; \quad \cos 2\theta_3 = \frac{2}{\sqrt{5}}. \quad (17)$$

3.2 Quantum Copying

Once the qubits of the quantum copier are properly prepared then the copying of the initial state $|\Psi\rangle_{a_0}^{(in)}$ of the original qubit can be performed by a sequence of four controlled-NOT operations

$$|\Psi\rangle_{a_0 a_1 b_1}^{(out)} = \hat{P}_{b_1 a_0} \hat{P}_{a_1 a_0} \hat{P}_{a_0 b_1} \hat{P}_{a_0 a_1} |\Psi\rangle_{a_0}^{(in)} |\Psi\rangle_{a_1 b_1}^{(prep)}. \quad (18)$$

When this operation is combined with the preparation stage, we find that the basis states of the original qubit (a_0) are copied as described by Eq. (6) with $|\uparrow\rangle_x \equiv |0\rangle_{b_1}$ and $|\downarrow\rangle_x \equiv |1\rangle_{b_1}$.

4 Multiple Copying

Here we present a generalization of the quantum network (18) to the case when a set of N copy qubits a_j ($j = 1, \dots, N$) are produced out of the original qubit a_0 [10].

To find the $1 \rightarrow 1 + N$ network we assume the following:

(1) We assume that the information from the original qubit is copied to N copy qubits a_j which are initially prepared in the state $|N; 0\rangle_{\mathbf{a}} \equiv |0\rangle_{a_1} \dots |0\rangle_{a_N}$ (here the subscript \mathbf{a} is a shorthand notation indicating that $|N; 0\rangle_{\mathbf{a}}$ is a vector in the Hilbert space of N qubits a_j).

(2) To implement multiple quantum copying we need to associate an *idle* qubit b_j with each copy qubit, a_j . These N idle qubits, which play the role of the copying machine itself, are initially prepared in the state $|N; 0\rangle_{\mathbf{b}} \equiv |0\rangle_{b_1} \dots |0\rangle_{b_N}$.

(3) Prior to the transfer of information from the original qubit, the copy and the idle qubits have to be prepared in a specific state $|\Psi\rangle_{\mathbf{ab}}^{(prep)}$

$$|\Psi\rangle_{\mathbf{ab}}^{(prep)} = \sum_{k=0}^N [e_k |N; k\rangle_{\mathbf{a}} + f_k |N; k-1\rangle_{\mathbf{a}}] |N; k\rangle_{\mathbf{b}}, \quad (19)$$

where

$$e_k = \sqrt{\frac{2}{N+2}} \frac{\binom{N}{k}}{\binom{N+1}{k}}; \quad f_k = \sqrt{\frac{k}{N-k+1}} e_k, \quad (20)$$

and $|N; k\rangle_{\mathbf{a}}$ are normalized *symmetric* N -qubit state vectors with k qubits in the state $|1\rangle$ and $(N-k)$ qubits in the state $|0\rangle$. The states (19) can be obtained by performing a sequence of local rotations \mathbf{R} and controlled-NOT operations analogous to Eq. (16) [15]. Once the copying machine is prepared in the state $|\Psi\rangle_{\mathbf{ab}}^{(prep)}$ we can start to copy information from the original qubit a_0 .

To describe the copying network we firstly introduce an operator $\hat{Q}_{a_0 \mathbf{a}}$ which is a product of the controlled-NOTs defined by Eq. (13) with a_0 being a control qubit and a_j ($j = 1, \dots, N$) being targets:

$$\hat{Q}_{a_0 \mathbf{a}} \equiv \hat{P}_{a_0 a_N} \hat{P}_{a_0 a_{N-1}} \dots \hat{P}_{a_0 a_1}. \quad (21)$$

We also introduce the operator $\hat{Q}_{\mathbf{a}a_0}$ describing the controlled-NOT process with a_0 playing the role of the target qubit, i.e.

$$\hat{Q}_{\mathbf{a}a_0} \equiv \hat{P}_{a_N a_0} \hat{P}_{a_{N-1} a_0} \dots \hat{P}_{a_1 a_0}. \quad (22)$$

Now we find the $1 \rightarrow 1 + N$ copying network to be

$$|\Psi\rangle_{a_0}^{(in)} |N; 0\rangle_{\mathbf{a}} |N; 0\rangle_{\mathbf{b}} \longrightarrow |\Psi\rangle_{a_0}^{(in)} |\Psi\rangle_{\mathbf{ab}}^{(prep)} \longrightarrow |\Psi\rangle_{a_0 \mathbf{ab}}^{(out)}, \quad (23)$$

where the $(2N + 1)$ qubit output of the copying process is described by the state vector $|\Psi\rangle_{a_0 \mathbf{ab}}^{(out)}$ which is obtained after the following sequence of operations

$$|\Psi\rangle_{a_0 \mathbf{ab}}^{(out)} = \hat{Q}_{\mathbf{b}a_0} \hat{Q}_{\mathbf{a}a_0} \hat{Q}_{a_0 \mathbf{b}} \hat{Q}_{a_0 \mathbf{a}} |\Psi\rangle_{a_0}^{(in)} \cdot |\Psi\rangle_{\mathbf{ab}}^{(prep)}. \quad (24)$$

This last equation describes a simple quantum network when firstly the original qubit controls the target qubits of the quantum copier. Then the qubits \mathbf{a} and \mathbf{b} “control” the state of the original qubit via another sequence of controlled-NOTs. In this way one can produce out of a single original qubit a set of quantum clones. This quantum network realizes the unitary transformation for $1 \rightarrow 1 + N$ cloning as introduced by Gisin and Massar [8].

5 Properties of Copied Qubits

Using the explicit expression for the output state $|\Psi\rangle_{a_0 \mathbf{ab}}^{(out)}$ we find that the original and the copy qubits at the output of the quantum copier are in the same state described by the density operator

$$\hat{\rho}_{a_j}^{(out)} = s^{(N)} \hat{\rho}_{a_j}^{(id)} + \frac{1 - s^{(N)}}{2} \hat{1}; \quad j = 0, 1, \dots, N, \quad (25)$$

where the scaling factor $s^{(N)}$ depends on the number N of copies, i.e.

$$s^{(N)} = \frac{1}{3} + \frac{2}{3(N+1)}, \quad (26)$$

which corresponds to the fidelity $\mathcal{F} = 2/3 + 1/3(N+1)$. We see that this result for $N = 1$ reduces to the case of the UQCM discussed in Section III. We also note that in the limit $N \rightarrow \infty$, i.e. when an infinite number of copies is *simultaneously* produced via the generalization of the UQCM, the copy qubits still carry information about the original qubit, because their density operators are given by the relation

$$\hat{\rho}_{a_j}^{(out)} = \frac{1}{3} \hat{\rho}_{a_j}^{(id)} + \frac{1}{3} \hat{1}; \quad j = 0, 1, \dots, \infty, \quad (27)$$

which corresponds to the fidelity $\mathcal{F} = 2/3$. This is the optimal fidelity achievable when an *optimal* measurement is performed on a single qubit [16, 17]. From this

point of view one can consider quantum copying as a transformation of quantum information into classical information [8]. This also suggests that quantum copying can be utilized to obtain novel insights into the quantum theory of measurement [e.g., a simultaneous measurement of conjugated observables on two copies of the original qubit; or a specific realization of the generalized (POVM) measurement performed on the original qubit].

Comment 1

We note that idle qubits b_j after the copying is performed are always in the state

$$\hat{\rho}_{b_j}^{(out)} = \frac{1}{3} \left(\hat{\rho}_{b_j}^{(id)} \right)^T + \frac{1}{3} \hat{1}, \quad j = 1, \dots, N, \quad (28)$$

irrespective of the number of copies created from the original qubit.

Comment 2

Using the Peres-Horodecki theorem [18,19] we can conclude that an arbitrary pair $\hat{\rho}_{a_m a_n}^{(out)}$ out of $N+1$ of copied qubits at the output of the copier is inseparable *only* in the case $N = 1$. In this case one of the eigenvalues of the partially transposed operator $[\hat{\rho}_{a_0 a_1}^{(out)}]^{T_2}$ is negative, which is the necessary and sufficient condition for the inseparability of the matrix $\hat{\rho}_{a_0 a_1}^{(out)}$. For $N > 1$ all pairs of copied qubits at the output of the quantum copier are separable.

6 Cloning of Quantum Registers

In what follows we will propose a copy machine which universally copies higher dimensional systems. We shall be particularly interested in how the quality of the copies scales with the dimensionality, M , of the system being copied. What we find is that the fidelity of the copies decreases with M , as expected, but, somewhat surprisingly, does not go to zero as M goes to infinity.

Let us consider a quantum system prepared in a pure state which is described by the vector

$$|\Phi\rangle_{a_0} = \sum_{i=1}^M \alpha_i |\Psi_i\rangle_{a_0} \quad (29)$$

in an M -dimensional Hilbert space spanned by M orthonormal basis vectors $|\Psi_i\rangle_{a_0}$ ($i = 1, \dots, N$). The complex amplitudes α_i are normalized to unity, i.e. $\sum |\alpha_i|^2 = 1$. In particular, one can consider $M = 2^m$ where m is the number of qubits in a given quantum register. One can generalize the no-cloning theorem which has been proven for spin-1/2 particles (qubits) by Wootters and Zurek [1] for arbitrary quantum systems. That is, there does not exist a unitary transformation such that the state given in Eq. (29) can be ideally cloned (copied), i.e. it is impossible to find a unitary transformation such that $|\Phi\rangle_{a_0} \longrightarrow |\Phi\rangle_{a_0} |\Phi\rangle_{a_1}$ for an arbitrary input states.

Following our previous discussion we can ask whether a universal cloning transformation exists which will generate two imperfect copies from the original

state, $|\Phi\rangle_{a_0}$. The quality of the cloning should not depend on the particular state (in the given Hilbert space) which is going to be copied. This input-state independence (invariance) of the cloning can be formally expressed as

$$\hat{\rho}_{a_j}^{(out)} = s\hat{\rho}_{a_j}^{(id)} + \frac{1-s}{M}\hat{1}, \quad (30)$$

where $\hat{\rho}_{a_j}^{(id)} = |\Phi\rangle_{a_0 a_0} \langle \Phi|$ is the density operator describing the original state which is going to be copied. This scaling form of Eq. (30) guarantees that the Bures distance (3) between the input and the output density operators is input-state independent.

The quantum copying machine we shall use is itself an M dimensional quantum system, and we shall let $|X_i\rangle_x$ ($i = 1, \dots, M$) be an orthonormal basis of the copying machine Hilbert space. This copier is initially prepared in a particular state $|X\rangle_x$. The action of the cloning transformation can be specified by a unitary transformation acting on basis vectors of the tensor product space of the original quantum system $|\Psi_i\rangle_{a_0}$, the copier, and an additional M -dimensional system which is to become the copy (which is initially prepared in a state $|0\rangle_{a_1}$). We have found the transformation of the basis vectors $|\Psi_i\rangle_{a_0}$

$$\begin{aligned} |\Psi_i\rangle_{a_0} |0\rangle_{a_1} |X\rangle_x \longrightarrow & c |\Psi_i\rangle_{a_0} |\Psi_i\rangle_{a_1} |X_i\rangle_x \\ & + d \sum_{j \neq i}^M (|\Psi_i\rangle_{a_0} |\Psi_j\rangle_{a_1} + |\Psi_j\rangle_{a_0} |\Psi_i\rangle_{a_1}) |X_j\rangle_x; \end{aligned} \quad (31)$$

(where $i = 1, \dots, M$) with the coefficients

$$c = \frac{2}{\sqrt{2(M+1)}}; \quad d = \frac{1}{\sqrt{2(M+1)}}, \quad (32)$$

such that the input-state independence of cloning is satisfied. That is, the clones have density operators given by Eq. (30) with the scaling factor

$$s = c^2 + (M-2)d^2 = \frac{(M+2)}{2(M+1)}. \quad (33)$$

If $M = 2$, then the transformation (31) reduces to the copying transformation for qubits given by Eq. (6). From earlier results of Gisin and Massar [8] the optimality of the transformation (31) for $M = 2$ directly follows. At the moment we are not able to prove rigorously that the cloning transformation (31) is *optimal* for arbitrary $M > 2$. Nevertheless, we have performed numerical tests which suggest that the cloning transformation (31) is optimal.

We note that the scaling factor (33), which describes the quality of the copy, is a decreasing function of M . This is not surprising, because a quantum state in a large dimensional space contains more quantum information than one in a small dimensional one (e. g. a state in a 4 dimensional space contains information about 2 qubits while a state in a 2 dimensional one describes only a single qubit), so that as M increases one is trying to copy more and more quantum information.

On the other hand, it is interesting to note that in the limit $M \rightarrow \infty$, i.e. in the case when the Hilbert space of the given quantum system is infinite dimensional (e.g. quantum-mechanical harmonic oscillator), the cloning can still be performed efficiently with the scaling factor equal to $1/2$.

In order to confirm that the quality of the copies which the copying transformation (31) produces is input-state independent (i.e., all states are cloned equally well) we evaluate the Bures distance (3). In our particular case we find, that the distance between $\hat{\rho}_{a_k}^{(out)}$ and $\hat{\rho}_{a_k}^{(id)}$ depends only on the dimension of the Hilbert space M , but not on the state which is cloned, i.e.

$$d_B(\hat{\rho}_{a_k}^{(out)}, \hat{\rho}_{a_k}^{(id)}) = \sqrt{2} \left(1 - \sqrt{\frac{M+3}{2(M+1)}} \right)^{1/2}. \quad (34)$$

The Bures distance given by Eq. (34) is maximal when states in the infinite-dimensional Hilbert space are cloned, and in that case we find

$$\lim_{M \rightarrow \infty} d_B(\hat{\rho}_{a_k}^{(out)}, \hat{\rho}_{a_k}^{(id)}) = \sqrt{2 - \sqrt{2}}. \quad (35)$$

This means that even for an infinite-dimensional system, reasonable cloning can be performed, which is reflected in the fact that the corresponding scaling factor s is equal to $1/2$.

6.1 Local vs. Nonlocal Cloning

Finally, we compare two methods of copying quantum registers. In particular, we shall consider cloning an entangled state of two qubits. We assume that the two qubits are prepared in the state

$$|\Phi\rangle_{a_0 b_0} = \alpha|00\rangle_{a_0 b_0} + \beta|11\rangle_{a_0 b_0}, \quad (36)$$

where, for simplicity, we have taken α and β to be real, and $\alpha^2 + \beta^2 = 1$. First, we shall consider the case in which each of the two qubits a_0 and b_0 is copied *locally* by two independent quantum copiers [13]. Each of these two copiers is described by the transformation (31) with $M = 2$. Next, we shall consider a *nonlocal* cloning of the two-qubit state (36) when this system is cloned via the unitary transformation (31) with $M = 4$, i.e. the cloner in this case acts non-locally on the two qubits. Our chief task will be to analyze how inseparability is cloned in these two scenarios, but we shall also examine the quality of the copies which are produced. From the Peres-Horodecki theorem it follows that the state (36) is inseparable for all values of α^2 such that $0 < \alpha^2 < 1$.

Firstly, let us suppose that the two original qubits a_0 and b_0 are cloned independently (locally) by two independent local cloners X_I and X_{II} , each described by the transformation (31) with $M = 2$. The cloner X_I (X_{II}) generates out of qubit a_0 (b_0) two qubits a_0 and a_1 (b_0 and b_1). After we perform trace over the two cloners we obtain a four-qubit density operator $\hat{\rho}_{a_0 a_1 b_0 b_1}^{(out)}$ which also describes two nonlocal two-qubit systems, i.e. $\hat{\rho}_{a_0 b_1}$ and $\hat{\rho}_{a_1 b_0}$. These two two-qubit

systems are the clones of the original two-qubit register (36). We note that these density matrices cannot be expressed in the scaled form (30), and that the quality of the copies depends on the input state. From the Peres-Horodecki theorem we immediately find that the density operators $\hat{\rho}_{a_0b_1}$ and $\hat{\rho}_{a_1b_0}$ are inseparable if

$$\frac{1}{2} - \frac{\sqrt{39}}{16} \leq \alpha^2 \leq \frac{1}{2} + \frac{\sqrt{39}}{16}. \quad (37)$$

This proves that for a restricted set of pure two-qubit states (36) satisfying the condition (37), it is possible to perform a *local* cloning such that the original inseparability of entangled pair of qubits is (partially) preserved.

Secondly, let us see what happens when we copy the entire two-qubit register at once. We would like to determine whether the set of original two-qubit states (36), which after the cloning exhibit inseparability, is larger (i.e., the restriction of the form given in Eq. (37) is weaker) than when a local cloning is performed. To do so, we introduce four basis vectors $|\Psi_1\rangle = |00\rangle$; $|\Psi_2\rangle = |01\rangle$; $|\Psi_3\rangle = |10\rangle$; and $|\Psi_4\rangle = |11\rangle$, so that the original two-qubit state in Eq. (36) is expressed as $|\Phi\rangle = \alpha|\Psi_1\rangle + \beta|\Psi_4\rangle$. The copying is now performed according to the transformation (31) with $M = 4$. We find that each of the two pairs of two-qubit copies at the output of the copier is described by the same density operator, i.e. $\hat{\rho}_{a_0b_1} = \hat{\rho}_{a_1b_0}$. Moreover, the fidelity of copying is input-state independent, and the quality of cloned registers is higher than that in the case of local cloning. Again, using the Peres-Horodecki theorem we find that the density operators $\hat{\rho}_{a_0b_1}$ and $\hat{\rho}_{a_1b_0}$ are inseparable if

$$\frac{1}{2} - \frac{\sqrt{2}}{3} \leq \alpha^2 \leq \frac{1}{2} + \frac{\sqrt{2}}{3}. \quad (38)$$

We conclude that quantum inseparability can be copied better (i.e. for much larger range of the parameter α) by using a nonlocal copier than when two local copiers are used.

7 Conclusions

We have presented the universal optimal quantum copying (cloning) machine which optimally clones a single original qubit to $N + 1$ qubits. We have found a simple quantum network which realizes this quantum copier. In addition we have presented a universal cloner for quantum registers. We have numerically tested the optimality of this cloner, but a rigorous proof has yet to be found.

Quantum copiers can be effectively utilized in various processes designed for the manipulation of quantum information. In particular, quantum copiers can be used for eavesdropping [20], they can be applied for realization of generalized (POVM) measurements [21], or they can be utilized for storage and retrieval of information in quantum computers [22].

Acknowledgments

We thank Sam Braunstein, Dagmar Bruß, Peter Knight, Martin Plenio, and Vlatko Vedral for collaboration and helpful discussions. This work was supported by the Royal Society.

References

1. W.K. Wootters and W.H. Zurek, *Nature* **299**, 802 (1982).
2. D. Diekes, *Phys. Lett. A* **92**, 271 (1982).
3. H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa, and B. Schumacher, *Phys. Rev. Lett.* **76**, 2818 (1996).
4. C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W.K. Wootters, *Phys. Rev. Lett.* **76**, 722 (1996).
5. V. Bužek and M. Hillery, *Phys. Rev. A* **54**, 1844 (1996).
6. V. Bužek and M. Hillery, *Acta Physica Slovaca* **47**, 193 (1997).
7. V. Bužek, S. Braunstein, M. Hillery, and D. Bruß, *Phys. Rev. A* **56**, 3446 (1997).
8. N. Gisin and S. Massar, *Phys. Rev. Lett.* **79**, 2153 (1997).
9. D. Bruß, D.P. DiVincenzo, A.K. Ekert, C. Machiavello, and J. Smolin: “Optimal universal and state-dependent quantum cloning,” [Los Alamos e-print archive quant-ph/9703046 (1997)].
10. V. Bužek, M. Hillery, and P.L. Knight, “Flocks of quantum clones: Multiple copying of qubits”, to appear in the special issue of *Fort. der Physik* edited by S. Braunstein.
11. M. Hillery and V. Bužek, *Phys. Rev. A* **56**, 1212 (1997).
12. D. Bruß, A.K. Ekert, and C. Machiavello, “Optimal universal quantum cloning and state estimation,” [Los Alamos e-print archive quant-ph/9712019(1997)].
13. V. Bužek, V. Vedral, M. Plenio, P.L. Knight and M. Hillery, *Phys. Rev. A* **55**, 3327 (1997).
14. D. Bures, *Trans. Am. Math. Soc.* **135**, 199 (1969); see also A. Uhlmann, *Rep. Math. Phys.* **9**, 273 (1976); **24** 229 (1986); W.K. Wootters, *Phys. Rev. D* **23**, 357 (1981).
15. A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J.A. Smolin, and H. Weinfurter, *Phys. Rev. A* **52**, 3457.
16. S. Massar and S. Popescu, *Phys. Rev. Lett.* **74**, 1259 (1995).
17. R. Derka, V. Bužek, and A.K. Ekert, *Phys. Rev. Lett.* **80**, 1571 (1998), and references therein.
18. A. Peres, *Phys. Rev. Lett.* **77**, 1423 (1996).
19. M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Lett. A* **223**, 1 (1997).
20. N. Gisin and B. Huttner, *Phys. Lett. A* **228**, 13 (1997).
21. R. Derka and V. Bužek: “Optimal estimation of quantum states from finite ensembles: From pure theory to hypothetic experiments” (unpublished).
22. D.P. DiVincenzo, *Science* **279**, 255 (1995).

Entanglement of Assistance

David P. DiVincenzo¹, Christopher A. Fuchs², Hideo Mabuchi²,
John A. Smolin¹, Ashish Thapliyal³, and Armin Uhlmann⁴

¹ IBM Research Division, Yorktown Heights, NY 10598, U. S. A.
divince/smolin@watson.ibm.com

² Bridge Laboratory of Physics 12-33, California Institute of Technology, Pasadena,
CA 91125, U. S. A. cfuchs/hmabuchi@cco.caltech.edu

³ Department of Physics, University of California at Santa Barbara, Santa Barbara,
CA 93106, U. S. A. ash@physics.ucsb.edu

⁴ Institut für Theoretische Physik, Universität Leipzig, Augustusplatz 10/11,
D-04109 Leipzig, Germany. Armin.Uhlmann@itp.uni-leipzig.de

Abstract. The newfound importance of “entanglement as a resource” in quantum computation and quantum communication behooves us to quantify it in as many distinct ways as possible. Here we explore a new quantification of entanglement of a general (mixed) quantum state for a bipartite system, which we name *entanglement of assistance*. We show it to be the maximum of the average entanglement over all ensembles consistent with the density matrix describing the bipartite state. With the help of lower and upper bounds we calculate entanglement of assistance for a few cases and use these results to show the surprising property of superadditivity. We believe that this may throw some light on the question of additivity of *entanglement of formation*.

1 Introduction

Much of the preoccupation of classical information theory is in making the correlation between two ends of a communication channel—that is, between a sender and a receiver—as high as possible. This is what (classical) communication is about. In contrast, this is only part of the story for the fledgling field of *quantum* information theory. The quantum world brings with it a new resource that senders and receivers can share: quantum entanglement, the stuff Einstein-Podolsky-Rosen pairs and Bell-inequality violations are made of. This new resource, of all the things in quantum information theory, is the most truly “quantum” of the lot. It is a resource because of the myriad ways in which it is starting to be used. The list of its applications already includes quantum key distribution [1], quantum-state teleportation [2], entanglement-enhanced classical communication [3], error-correction for quantum computers [4], entanglement-assisted multi-party communication [5], and better control of atomic frequency standards [6] among other things. The list grows every day.

The newfound importance of “entanglement as a resource” behooves us to quantify entanglement in as many distinct ways as possible. To this end, we

explore a new quantification of the entanglement of a general (mixed) quantum state for a bipartite system. This quantity we dub the *entanglement of assistance*. It is something of a dual notion to that of the *entanglement of formation* studied by Bennett, DiVincenzo, Smolin and Wootters [7]. It can be motivated in the following way.

Consider three players Alice, Bob, and Charlie, who jointly possess many copies of a tripartite quantum system each described by the pure state $|\Psi^{ABC}\rangle$. It follows that Alice and Bob, considered in isolation, possess many copies of the (generally mixed) state $\rho^{AB} = \text{tr}_C |\Psi^{ABC}\rangle\langle\Psi^{ABC}|$. Suppose now that Alice and Bob need to use their shared quantum system for one of the tasks above, teleportation for instance. However, very unfortunately, Charlie is not immediately available to pass his systems over for their use. Since their state may not be very pure or entangled for that matter, they may be seriously impaired by this. How might Charlie help in spite of his constraint?

One thing he can do is perform a measurement on each system that he possesses so that Alice and Bob will have pure states conditioned on his outcomes. If he transmits the classical information so obtained, they may be able to use it to their benefit. For instance, if several of his measurements reveal that Alice and Bob actually possess the pure state $|\Psi^{AB}\rangle$, then by the result of Bennett, et al. [8], they can convert these quantum states into maximally-entangled singlet states at a rate of $S(\rho^A) = -\text{tr}(\rho^A \log_2 \rho^A)$, where $\rho^A = \text{tr}_B |\Psi^{AB}\rangle\langle\Psi^{AB}|$. In general then, for a measurement that creates the states $|\Psi_i^{AB}\rangle$ with probabilities p_i , i.e., an ensemble $\mathcal{E} = \{p_i, |\Psi_i^{AB}\rangle\langle\Psi_i^{AB}|\}$, the average rate for conversion to singlets will be

$$E(\mathcal{E}) = \sum_i p_i S(\rho_i^A) = - \sum_i p_i \text{tr}(\rho_i^A \log_2 \rho_i^A) . \quad (1)$$

Since Charlie is a friend of Alice and Bob, he should choose his measurement so that this average rate is maximized. By a theorem of Hughston, Jozsa, and Wootters [9] this maximization can be taken over all possible ensembles \mathcal{E} consistent with ρ^{AB} in the sense that

$$\rho^{AB} = \sum_i p_i |\Psi_i^{AB}\rangle\langle\Psi_i^{AB}| .$$

Therefore the resulting maximal rate is a number intrinsic to ρ^{AB} , depending upon no further details of the state $|\Psi^{ABC}\rangle$. This is the *entanglement of assistance* $A(\rho^{AB})$ for ρ^{AB} :

$$A(\rho^{AB}) = \max_{\mathcal{E}} E(\mathcal{E}) . \quad (2)$$

This quantity seems to be dual to the entanglement of formation, which is defined by

$$F(\rho^{AB}) = \min_{\mathcal{E}} E(\mathcal{E}) . \quad (3)$$

In this paper, we demonstrate several properties possessed by entanglement of assistance, calculating it explicitly for some specialized cases. We exhibit a general upper bound, a particular upper bound specialized to the two-qubit case and a lower bound for diagonal density matrices, on the entanglement of

assistance. With the aid of the last two, we prove a fairly surprising property: the entanglement of assistance is in some cases superadditive. That is, there exist density operators $\rho_{1,2}^{AB}$ for which

$$A(\rho_1^{AB} \otimes \rho_2^{AB}) > A(\rho_1^{AB}) + A(\rho_2^{AB}) . \quad (4)$$

This means that if Charlie performs entangled measurements on his separate copies of the system, he can be more effective in helping Alice and Bob than he would have been otherwise.

2 Properties and Bounds

The average entanglement of an ensemble is invariant under local unitary transformations and non-increasing under general local operations¹ with classical communication [7,8]; so entanglement of assistance also has these properties. In addition to this it is easy to see that since there is a maximization over all possible pure-state realizations of the density matrix, the entanglement of assistance A is concave. That is,

$$A\left(\sum_i p_i \rho_i\right) \geq \sum_i p_i A(\rho_i) . \quad (5)$$

We note here that the entanglement of assistance is the least of all concave functions coinciding on pure states with their partial entropy i.e. entropy as seen from one side, while entanglement of formation is the greatest convex function with that property. This may be seen as a sign of the mentioned duality between the two.

Since we do not have a general formula² for entanglement of assistance of a given density matrix, upper and lower bounds are of great importance. We have found a few bounds which have helped us calculate the entanglement of assistance in some cases and have been of immense value in proving its superadditivity.

2.1 Upper Bounds

We have found an upper bound — the entropic bound — which works in general, and another upper bound — the fidelity bound — which works for the simplest case of one qubit on Alice's side and one on Bob's side.

¹ Local operations are those that Alice and Bob can perform on their own part of the system with possible synchronization using classical messages. However, they may not send each other quantum systems nor can they come together and perform a joint quantum operation on the system.

² We will present a formula for a non-trivial class of rank two density matrices in a forthcoming publication.

Entropic Bound: The entanglement of assistance is never greater than the minimum of the partial entropy as seen by Alice or Bob:

$$A(\rho) \leq \min[S(\text{tr}_A \rho), S(\text{tr}_B \rho)] \quad (6)$$

To prove this assume that $A(\rho)$ is achieved by the pure-state ensemble $\mathcal{E} = \{p_i, \Pi_i = |\psi_i\rangle\langle\psi_i|\}$, then by the concavity of Von Neumann entropy S we have,

$$S(\text{tr}_A \rho) \geq \sum_i p_i S(\text{tr}_A \Pi_i) = A(\rho) \quad (7)$$

The same inequality holds with trace over A replaced by trace over B. Combining them we get the entropic bound.

Since the Von Neumann entropy is additive and concave the bound is also additive and concave. Thus for states which saturate this bound the entanglement of assistance must be additive. We can also see from the results of section (3) that the entropic bound is zero if and only if the entanglement of assistance is zero. The bound obviously agrees with entanglement of assistance for pure states. It is easy to show³ that for the case of one qubit each on Alice's and Bob's side, the entropic bound is maximum (1 ebit) if and only if the entanglement of assistance is also maximum.

Fidelity Bound: For the case of one qubit each with Alice and Bob, we can get a bound that is sometimes stronger than the entropic bound. It says that the entanglement of assistance cannot be greater than the fidelity of the density matrix relative to its complex conjugate in the magic basis⁴ i.e.,

$$A(\rho) \leq F(\rho, \tilde{\rho}) \quad (8)$$

where the $\tilde{}$ — Hill-Wootters tilde — represents complex conjugation in the magic basis [10] and $F(\rho, \sigma) = \text{tr} \sqrt{\rho^{1/2} \sigma \rho^{1/2}}$ is the fidelity [11] that is, the square root of the transition probability [12]. To prove the bound the main thing we need is a classical inequality on the binary Shannon entropy [13]. It is,

$$H_2(x) = -x \log_2 x - (1-x) \log_2 (1-x) \quad (9)$$

$$\leq 2\sqrt{x(1-x)} \quad (10)$$

Equality is achieved in this if and only if $x = 0, \frac{1}{2}, 1$. The other facts we need are that for a pure state Π_i in a decomposition of ρ , the largest eigenvalue, $\lambda_1(\rho_i^A = \text{tr}_B \Pi_i)$, of its reduced density operator is given in terms of the Hill-Wootters tilde by

$$\lambda_1(\rho_i^A) = \frac{1}{2} \left[1 + \sqrt{1 - \text{tr}(\Pi_i \tilde{\Pi}_i)} \right] \quad (11)$$

³ We show this in a forthcoming publication.

⁴ The magic basis is:

$$|e_1\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}; |e_2\rangle = \frac{|00\rangle - |11\rangle}{-i\sqrt{2}}; |e_3\rangle = \frac{|01\rangle + |10\rangle}{-i\sqrt{2}}; |e_4\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

and, that the fidelity function is doubly concave in its two arguments and

$$F(\Pi_i, \tilde{\Pi}_i) = \sqrt{\text{tr}(\Pi_i \tilde{\Pi}_i)} . \quad (12)$$

Putting all this together: for any ensemble $\mathcal{E} = \{p_i, \Pi_i\}$ consistent with ρ average entanglement is

$$E(\mathcal{E}) = \sum_i p_i S(\rho_i^A) \quad (13)$$

$$= \sum_i p_i H_2(\lambda_1(\rho_i^A)) \quad (14)$$

$$\leq 2 \sum_i p_i \sqrt{\lambda_1(\rho_i^A)(1 - \lambda_1(\rho_i^A))} \quad (15)$$

$$= \sum_i p_i \sqrt{\text{tr}(\Pi_i \tilde{\Pi}_i)} \quad (16)$$

$$= \sum_i p_i F(\Pi_i, \tilde{\Pi}_i) \quad (17)$$

$$\leq F(\rho, \tilde{\rho}) . \quad (18)$$

The bound in Eq. (8) follows immediately.

Note the conditions for saturating this bound. In particular, in going from Eq. (14) to (15), we see that this new bound can be achieved only if ρ has a decomposition consisting only of maximally entangled and completely unentangled states. Thus the bound can't be tight generally—consider simply a ρ that is pure but not itself maximally entangled. However, we will see in section (4) that this bound is sometimes better than the entropic bound.

2.2 Lower Bounds

To get lower bounds we use the fact that the average entanglement of any pure-state realization of the density matrix gives us a lower bound on the entanglement of assistance. Using this idea we can come up with lower bounds by making systematic pure-state decompositions of the density matrix. For example, average entanglement of the eigenvector decomposition of the density matrix is a lower bound.

Applying the above idea to density matrices diagonal in a product basis we get a lower bound which we name *diagonal lower bound*. Consider the case of only one qubit on either side so that the product basis can be written as $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. The density matrix in this product basis looks like,

$$\rho = \begin{pmatrix} \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & 0 & \alpha_4 \end{pmatrix} \quad (19)$$

and the bound is,

$$A(\rho) \geq (\alpha_1 + \alpha_4)H_2\left(\frac{\alpha_1}{\alpha_1 + \alpha_4}\right) + (\alpha_2 + \alpha_3)H_2\left(\frac{\alpha_2}{\alpha_2 + \alpha_3}\right) . \quad (20)$$

This follows from the pure state realization of ρ : $\mathcal{E} = \{p_i, |\psi_i\rangle \mid i = 1, \dots, 4\}$, with

$$\begin{aligned} \{p_i\} &= \{(\alpha_1 + \alpha_4)/2, (\alpha_1 + \alpha_4)/2, (\alpha_2 + \alpha_3)/2, (\alpha_2 + \alpha_3)/2\} \\ |\psi_1\rangle &= (\sqrt{\alpha_1/(\alpha_1 + \alpha_4)}, 0, 0, \sqrt{\alpha_4/(\alpha_1 + \alpha_4)}) \\ |\psi_2\rangle &= (\sqrt{\alpha_1/(\alpha_1 + \alpha_4)}, 0, 0, -\sqrt{\alpha_4/(\alpha_1 + \alpha_4)}) \\ |\psi_3\rangle &= (0, \sqrt{\alpha_2/(\alpha_2 + \alpha_3)}, \sqrt{\alpha_3/(\alpha_2 + \alpha_3)}, 0) \\ |\psi_4\rangle &= (0, \sqrt{\alpha_2/(\alpha_2 + \alpha_3)}, -\sqrt{\alpha_3/(\alpha_2 + \alpha_3)}, 0) . \end{aligned} \quad (21)$$

Numerical results using optimization algorithms suggest that this is indeed the exact formula for entanglement of assistance for these density matrices. Also when this result saturates any of the upper bounds we know that it is the answer.

The bound can easily be generalized to higher dimensions⁵. Another thing to note here is that off-diagonal terms between $\{|00\rangle, |11\rangle\}$ and $\{|01\rangle, |10\rangle\}$ don't change the lower bound because we can still use a similar pure state decomposition as before, with slightly modified probabilities.

3 Examples

Let us start by characterizing the states that have zero entanglement of assistance. We find that there is a very simple characterization of these states as, states whose density matrix is pure on at least one side. That is, either $\rho = |\psi_A\rangle\langle\psi_A| \otimes \rho_B$ or $\rho = \rho_A \otimes |\psi_B\rangle\langle\psi_B|$.

To prove this result first note that since the entropic (upper) bound (6) is zero for such states, these states have zero entanglement of assistance. Let us now prove that only such states can have zero entanglement of assistance. Since entanglement of assistance is zero, every valid pure-state decomposition of ρ must consist exclusively of product states. (For otherwise the average entanglement of that ensemble will be necessarily greater than zero.) So let us focus on an eigendecomposition of ρ : this is an ensemble $\mathcal{E} = \{\lambda_i, |\alpha_i\rangle|\beta_i\rangle\}$. (The $\lambda_i \neq 0$ are the probabilities for the associated states.)

Now, by the Hughston-Jozsa-Wootters theorem [9], any “unitary reshuffling” of this ensemble is also a valid ensemble. Therefore let us consider a new ensemble $\mathcal{E}' = \{p_i, |\psi_i\rangle\}$ that is exactly the same as the old one, except in the first two elements. Namely,

$$\sqrt{p_1}|\psi_1\rangle = \cos\theta\sqrt{\lambda_1}|\alpha_1\rangle|\beta_1\rangle + \sin\theta\sqrt{\lambda_2}|\alpha_2\rangle|\beta_2\rangle \quad (22)$$

$$\sqrt{p_2}|\psi_2\rangle = \sin\theta\sqrt{\lambda_1}|\alpha_1\rangle|\beta_1\rangle - \cos\theta\sqrt{\lambda_2}|\alpha_2\rangle|\beta_2\rangle . \quad (23)$$

⁵ This will be presented in a forthcoming publication.

Because the ensemble \mathcal{E} is an eigenensemble, we must have either $\langle \alpha_1 | \alpha_2 \rangle = 0$ or $\langle \beta_1 | \beta_2 \rangle = 0$ or both. First note that it cannot be the case that both inner products vanish. For otherwise, there will be at least one value of θ for which $|\psi_1\rangle$ and $|\psi_2\rangle$ will be entangled states. Consequently, let us look at one of the other possibilities: namely, $|\alpha_1\rangle$ and $|\alpha_2\rangle$ are orthogonal, but $\langle \beta_1 | \beta_2 \rangle \neq 0$. Then we must have $|\beta_1\rangle = e^{i\phi} |\beta_2\rangle$. For otherwise, again—for at least one value of θ —both $|\psi_1\rangle$ and $|\psi_2\rangle$ will be entangled states. The argument goes analogously if $|\beta_1\rangle$ and $|\beta_2\rangle$ are orthogonal. That is to say, $|\psi_1\rangle$ and $|\psi_2\rangle$ must indeed be product states but with either identical left factors or identical right factors.

Therefore, just to say it again (but now ignoring phases, since they could have been in the “unitary reshuffling” anyway), we must have either $|\alpha_1\rangle = |\alpha_2\rangle$ or $|\beta_1\rangle = |\beta_2\rangle$. Locking one of these possibilities in, we can go through exactly the same argument for similar “unitary reshufflings” of other pairs of eigenstates. Consistency then gives that we must find either

$$\rho = |\psi_A\rangle\langle\psi_A| \otimes \rho_B \quad \text{or} \quad \rho = \rho_A \otimes |\psi_B\rangle\langle\psi_B|. \quad (24)$$

This completes the proof.

Let us now look at the states with the maximum value [\[6\]](#) for entanglement of assistance. This is the set of all possible mixtures of maximally entangled pure-states. For the simplest non-trivial case of one qubit on either side, this is equivalent to the set of ‘T’ states of Horodecki et al. [\[14,15\]](#) or equivalently the set of real [\[7\]](#) density matrices in the magic basis [\[10\]](#).

Next we consider some non-extremal examples. Consider states like $\rho = \text{Diagonal}[\alpha, 0, 0, 1 - \alpha]$ in a product basis. The entropic bound [\(6\)](#) says $A(\rho) \leq H_2(\alpha)$ and the diagonal lower bound [\(20\)](#) gives us $A(\rho) \geq H_2(\alpha)$ so that $A(\rho) = H_2(\alpha)$.

Let us look at a more complicated example, say $\rho = \text{Diagonal}[1/3, 1/3, 1/3, 0]$ in a product basis. The diagonal bound [\(20\)](#) gives us $A(\rho) \geq 2/3$. For the fidelity bound, we first calculate $\tilde{\rho} = \text{Diagonal}[0, 1/3, 1/3, 1/3]$ and by the fidelity bound [\(8\)](#) we see that $A(\rho) \leq 2/3$. Thus $A(\rho) = 2/3$. Notice that the entropic bound for this case is just $H_2(1/3) \approx 0.918$ ebits. Since it does not saturate the entropic bound this state is a good candidate for superadditivity. Let us turn to that next.

4 Superadditivity

It has been speculated (and is indicated numerically for a few cases) that the entanglement of formation is additive [\[16\]](#). This property is very important if it is to be used as a measure of entanglement. Since entanglement of assistance looks dual to entanglement of formation it comes as a bit of a surprise then that the entanglement of formation is superadditive. By this we mean

$$A(\rho_1^{\text{AB}} \otimes \rho_2^{\text{AB}}) > A(\rho_1^{\text{AB}}) + A(\rho_2^{\text{AB}}) \quad (25)$$

⁶ The maximum value is $\text{Log}_2(\text{Min}[N_A, N_B])$ ebits. Here, N_A and N_B are the dimensions of the Hilbert spaces on Alice’s and Bob’s side respectively.

⁷ upto a phase.

for some density matrices ρ_1^{AB} and ρ_2^{AB} . Since the entropic bound is additive, superadditivity is possible only for the states which do not saturate the bound.

The density matrix $\rho = \text{Diagonal}[1/3, 1/3, 1/3, 0]$ which we have just seen in section (B) has an entanglement of assistance $A = 2/3$. Recall that this state does not saturate the (additive) entropic bound. So it can show superadditivity and in fact it does! If A were additive: $A(\rho \otimes \rho) = 2A(\rho) = 4/3$. However $\rho \otimes \rho$ can be realized as an ensemble \mathcal{E} which has an average entanglement $E \approx 1.5506$. (This ensemble is shown in detail in appendix A). This being a lower bound on $A(\rho \otimes \rho)$ it shows that entanglement of assistance is superadditive.

Superadditivity of entanglement of assistance means that if Alice, Bob, and Charlie have two copies of the same system then Charlie can give Alice and Bob more entanglement if he makes an entangled measurement on the two copies of the system.

Since entanglement of assistance and entanglement of formation are dual to each other in the sense of replacing a maximization by a minimization, we would expect the superadditivity of the entanglement of assistance to tell us something about the additivity of the entanglement of formation. We have not been able to see a connection as yet but we note here a result [17],

$$A(\rho) - F(\rho) \leq S(\rho) - |S(\text{tr}_A \rho) - S(\text{tr}_B \rho)|, \quad (26)$$

which connects entanglement of assistance and entanglement of formation. Thus in addition to its own physical significance, entanglement of assistance may turn out to be a useful tool to study the problem of additivity of the entanglement of formation.

5 Conclusions

In this paper we have introduced a new measure of entanglement — entanglement of assistance — and studied its properties and found upper and lower bounds for it. We have proved that it is superadditive meaning that, entangled measurements by Charlie give more entanglement to Alice and Bob. We also note here that the superadditivity of entanglement of assistance may throw some light upon the additivity question of the entanglement of formation.

6 Acknowledgments

Part of this work was completed during the 1997 Elsag-Bailey – I.S.I. Foundation research meeting on quantum computation. We would like to thank Charles H. Bennet, William K. Wootters and Barbara M. Terhal for illuminating discussions. DPD and JAS would like to thank the Army Research Office for support. CAF was supported by a Lee A. DuBridge Fellowship and by DARPA through the Quantum Information and Computing (QUIC) Institute administered by the US Army Research Office. AT would like to thank The NSF Science and Technology Center for Quantized Electronics Structures, Grant #DMR 91-20007, for

support to attend this conference and, IBM Research Division for supporting his summer visit during which this work was initiated. He would also like to thank Prof. David Awschalom for his invaluable support, without which, it would have been impossible to work in this exciting field.

A Superadditivity

Here we look at the ensemble which gives us an average entanglement more than the additive value of $4/3$ for the density matrix discussed in section (4). The ensemble written in the product basis $\{|00_A00_B\rangle, |00_A01_B\rangle, \dots, |11_A11_B\rangle\}$ is, $\mathcal{E} = \{p_i, |\psi_i\rangle, i = 1, \dots, 12\}$ where,

$$\begin{aligned}
 \{p_i\} &= (\alpha_1, \alpha_1, \alpha_1, \alpha_1, \alpha_1, \alpha_1, \alpha_2, \alpha_2, \alpha_2, \alpha_2, \alpha_2, \alpha_2) \\
 |\psi_1\rangle &= (0, 0, 0, a, 0, 0, be^{i\pi/3}, 0, 0, be^{-i\pi/3}, 0, 0, a, 0, 0, 0) \\
 |\psi_2\rangle &= (0, 0, 0, a, 0, 0, -be^{i\pi/3}, 0, 0, -be^{-i\pi/3}, 0, 0, a, 0, 0, 0) \\
 |\psi_3\rangle &= (0, 0, 0, a, 0, 0, be^{i2\pi/3}, 0, 0, be^{-i2\pi/3}, 0, 0, -a, 0, 0, 0) \\
 |\psi_4\rangle &= (0, 0, 0, a, 0, 0, -be^{i2\pi/3}, 0, 0, -be^{-i2\pi/3}, 0, 0, -a, 0, 0, 0) \\
 |\psi_5\rangle &= (d, 0, 0, 0, 0, 0, b, 0, 0, b, 0, 0, 0, 0, 0, 0) \\
 |\psi_6\rangle &= (d, 0, 0, 0, 0, 0, -b, 0, 0, -b, 0, 0, 0, 0, 0, 0) \\
 |\psi_7\rangle &= (0, c, 0, 0, 0, 0, b, 0, c, 0, 0, 0, 0, 0, 0, 0) \\
 |\psi_8\rangle &= (0, c, 0, 0, 0, 0, be^{i2\pi/3}, 0, ce^{-i2\pi/3}, 0, 0, 0, 0, 0, 0, 0) \\
 |\psi_9\rangle &= (0, c, 0, 0, 0, 0, be^{i4\pi/3}, 0, ce^{-i4\pi/3}, 0, 0, 0, 0, 0, 0, 0) \\
 |\psi_{10}\rangle &= (0, 0, c, 0, c, 0, 0, 0, 0, b, 0, 0, 0, 0, 0, 0) \\
 |\psi_{11}\rangle &= (0, 0, c, 0, ce^{i2\pi/3}, 0, 0, 0, 0, be^{-i2\pi/3}, 0, 0, 0, 0, 0, 0) \\
 |\psi_{12}\rangle &= (0, 0, c, 0, ce^{i4\pi/3}, 0, 0, 0, 0, be^{-i4\pi/3}, 0, 0, 0, 0, 0, 0)
 \end{aligned}$$

and,

$$\begin{aligned}
 a &= \left(6 - \frac{24}{5 + \sqrt{7}}\right)^{-\frac{1}{2}} \approx 0.5912 \\
 b &= \left(9 - \frac{18}{5 + \sqrt{7}}\right)^{-\frac{1}{2}} \approx 0.3879 \\
 c &= \left(\frac{5 + \sqrt{7}}{18}\right)^{\frac{1}{2}} \approx 0.6517 \\
 d &= \left(3 - \frac{12}{5 + \sqrt{7}}\right)^{-\frac{1}{2}} \approx 0.8361 \\
 \alpha_1 &= \frac{1}{6} - \frac{2}{3(5 + \sqrt{7})} \approx 0.0795 \\
 \alpha_2 &= \frac{2}{3(5 + \sqrt{7})} \approx 0.0872.
 \end{aligned} \tag{27}$$

The entanglements for the pure-states forming the ensemble \mathcal{E} are,

$$\{E_i\} = (E_\alpha, E_\alpha, E_\alpha, E_\alpha, E_\beta, E_\beta, E_\gamma, E_\gamma, E_\gamma, E_\gamma, E_\gamma, E_\gamma) \quad (28)$$

where,

$$\begin{aligned} E_\alpha &= H_2(a^2, a^2, b^2, b^2) \approx 1.8824 \text{ ebits} , \\ E_\beta &= H_2(d^2, b^2, b^2, 0) \approx 1.1834 \text{ ebits} , \\ E_\gamma &= H_2(c^2, b^2, c^2, 0) \approx 1.4605 \text{ ebits} , \end{aligned}$$

and average entanglement of this ensemble is,

$$E(\mathcal{E}) = 4\alpha_1 E_\alpha + 2\alpha_1 E_\beta + 6\alpha_2 E_\gamma \approx 1.5506 \text{ ebits.} \quad (29)$$

To find this ensemble we use the following procedure: First, we use a numerical optimization program to find the ensemble which has the maximum average entanglement. Then using the structure of this ensemble, guessing the right phases and using the constraint that the ensemble must give us $\rho \otimes \rho$ we get a quadratic equation which we solve to find the values for a, b, c, d, α_1 and α_2 .

Note that according to our numerical results, the ensemble presented here achieves the maximum average entanglement consistent with the density matrix. This suggests that this is the entanglement of assistance for it.

References

1. C.H. Bennett and G. Brassard "Quantum Cryptography: Public Key Distribution and Coin Tossing", Proceedings of IEEE International Conference on Computers Systems and Signal Processing, Bangalore India, December 1984, pp 175-179.; D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).
2. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
3. C. H. Bennett, C. A. Fuchs, and J. A. Smolin, "Entanglement-Enhanced Classical Communication on a Noisy Quantum Channel," in *Quantum Communication, Computing and Measurement*, edited by O. Hirota, A. S. Holevo, and C. M. Caves (Plenum, New York, 1997).
4. P. W. Shor, Phys. Rev. **A 52**, 2493 (1995); D. Gottesman, Ph. D. Thesis, California Institute of Technology, 1997, LANL e-print [quant-ph/9705052](#).
5. R. Cleve and H. Buhrman, LANL e-print [quant-ph/9704026](#).
6. J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys. Rev. A **54**, R4649 (1996).
7. C. H. Bennet, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters, Phys. Rev. A **54**, 3824(1996), LANL e-print [quant-ph/9604024](#).
8. C. H. Bennett, H. J. Bernstein, S. Popescu, B. Schumacher, Phys. Rev. **A 53**, 2046(1996), LANL e-print [quant-ph/9511030](#).
9. L. P. Hughston, R. Jozsa, W. K. Wootters, Phys. Lett. A **183**, 14 (1993).

10. S. Hill, W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997), LANL e-print [quant-ph/9703041](#).
11. R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
12. A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
13. C. A. Fuchs and J. van de Graaf, LANL e-print [quant-ph/9712042](#).
14. R. Horodecki, M. Horodecki, Phys. Rev. A **54**, 1838 (1996).
15. R. Horodecki, M. Horodecki, P. Horodecki, Phys. Lett. A **222**, 21 (1996).
16. J. A. Smolin, private communication; W. K. Wootters, to appear in Phys. Rev. Lett., LANL e-print [quant-ph/9709029](#).
17. M. Nielsen, private communication; The same result was independently conjectured in a discussion with C. H. Bennet and W. K. Wootters.

What Information Theory Can Tell Us About Quantum Reality

C. Adami and N.J. Cerf

W. K. Kellogg Radiation Laboratory
California Institute of Technology, Pasadena, California 91125, USA

Abstract. An investigation of Einstein’s “physical” reality and the concept of quantum reality in terms of information theory suggests a solution to quantum paradoxes such as the Einstein-Podolsky-Rosen (EPR) and the Schrödinger-cat paradoxes. Quantum reality, the picture based on unitarily evolving wavefunctions, is complete, but appears incomplete from the observer’s point of view for fundamental reasons arising from the quantum information theory of measurement. Physical reality, the picture based on classically accessible observables is, in the worst case of EPR experiments, unrelated to the quantum reality it purports to reflect. Thus, quantum information theory implies that only correlations, not the correlata, are physically accessible: the mantra of the Ithaca interpretation of quantum mechanics.

1 Introduction

The concept of “physical reality” as championed by Einstein [1]—the postulate that the *objective* state of a system is specified by a set of real-valued parameters *independently* of our knowledge of them—has been an object of contention ever since the inception of quantum theory (see, e.g., [2,3,4,5,6,7]). The most prevailing views assert either that the “quantum reality” suggested by wavefunctions and non-local correlations is only a mathematical construction necessary for a consistent theory (Bohr’s view), or else that physical reality is deterministic but incompletely described by quantum mechanics (Einstein’s view). A popular interpretation of the latter view is that physical reality is obscured by inaccessible hidden variables [8], a stance that appears to be discredited by the violation of Bell’s inequalities in quantum mechanics [9]. Bohr’s view of complementarity, on the other hand, assigns a special status to classical physics as an essential ingredient in measurement since it requires the measurement device to be classical. As recognized by von Neumann [10], this undermines the foundations of quantum mechanics as a complete and consistent theory. Here, we suggest that Einstein realism and Bohr’s complementarity principle can be reconciled within a framework that consistently describes the concept of information in quantum mechanics. This is exemplified by the quantum information theoretic treatment of the Einstein-Podolsky-Rosen (EPR) experiment [11] and the Schrödinger-cat paradox [12], which has recently attracted increasing attention (see, e.g., [12]). We propose that, in general, the perceived physical reality and quantum reality can

be *disjoint*, that is, the result of a quantum measurement conceivably might not carry any information—in the sense of Shannon theory [13]—which would allow the observer to infer the state of the measured system. While counterintuitive, we shall show that this picture is a direct consequence of an information-theoretic reinterpretation of quantum measurement. Moreover, such a view effortlessly resolves the EPR paradox which has inspired the discussions on reality, as well as other quantum paradoxes rooted in the measurement problem.

The gedankenexperiment that constitutes the EPR paradox was created by Einstein, Podolsky, and Rosen to demonstrate their dissatisfaction with “unknowables” [1]. In that experiment, it appears that two complementary variables (such as position and momentum) are *in principle* measurable by exploiting the quantum correlations between the two particles, in contradiction with Heisenberg’s uncertainty principle. Their conclusion, namely that the quantum mechanical description of reality must therefore be incomplete, was based on a criterion for reality which they considered “reasonable” (see below). This criterion was faulted by Bohr [14] in his reply to the EPR paper, insisting rather that physical variables are never independent of the way they are measured owing to the complementarity principle, and therefore that measurements do not confer reality to properties of quantum objects. We shall show here, using quantum information theory only, that, while indeed an element of reality is *not* created for the measured quantum system, the result of a quantum measurement creates an element of reality for the result of *another* measurement, i.e., it allows you to predict the state of another measurement *device* without revealing the state of the quantum system itself. Thus, physical reality is reflected in *correlations* between classical objects only. This view, which we arrived at from a quantum information-theoretic examination of quantum measurement [15,16] essentially coincides with Mermin’s “Ithaca Interpretation of Quantum Mechanics”, Ref. [17].

2 The EPR Paradox

The EPR experiment in the version of Bohm [18] involves the preparation of a quantum system such as the one created by the decay of a spinless particle into two half-integral-spin particles:

$$|\Psi_{\text{EPR}}\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) . \quad (1)$$

This state represents the *superposition* of the two possible situations: “left-particle spin-up, right-particle spin-down”, and “right-particle spin-up, left-particle spin-down”. Let us now imagine that the pair so-created is separated sufficiently far that classical information would take a long time to travel between them. Then, we measure for example the *z*-component of the spin of one of the particles (say, the left one). This measurement has two possible outcomes, which occur with probability one-half each, implying that the von Neumann uncertainty of the density matrix describing any one of the particles (denoted by subscripts *L* and *R*),

$$\rho_{L,R} = \frac{1}{2} | \uparrow \rangle \langle \uparrow | + \frac{1}{2} | \downarrow \rangle \langle \downarrow | \quad (2)$$

is one bit

$$S(\rho_{L,R}) = -\text{Tr}_{R,L} (\rho_{L,R} \log_2 \rho_{L,R}) = 1 \quad (3)$$

in spite of the fact that entropy of the combined system *vanishes*. The latter is of course well-known: for a quantum mechanical “pure state” ($\rho_{\text{EPR}}^2 = \rho_{\text{EPR}}$, where $\rho_{\text{EPR}} = |\Psi_{\text{EPR}}\rangle\langle\Psi_{\text{EPR}}|$) the von-Neumann entropy vanishes $S(\rho_{\text{EPR}}) = 0$, i.e., the state is perfectly well-known.

Clearly then, the quantum nature of the EPR state is very peculiar since the uncertainty of a part of this system can be larger than the uncertainty of the pair. Classically, this is impossible. Indeed, if we describe uncertainties using (classical) Shannon entropies, the Shannon entropy of a system A , say, with $A \subset AB$, is

$$H(A) \leq H(AB) . \quad (4)$$

This property of *monotonicity* of entropies is violated in quantum mechanics [19]. This violation, on the other hand, can be described consistently in an information-theoretic formalism which allows for *negative* conditional entropies [20,21]. In other words, there exists an information theory, extended to the quantum regime, in which the violation of classical laws such as monotonicity are inevitable consequences.

Quantum *entanglement* situations, such as encountered in EPR pairs, are prototype systems to examine the classically forbidden regime of negative entropies. In the case at hand, the joint, conditional, mutual, and marginal entropies of the EPR pair can be summarized by the entropy diagram in Fig. 1. Such diagrams are used extensively in classical information theory and serve as mental scratch pads to remind us of the separation of unconditional entropies into conditional and mutual pieces. While in the past the violation of monotonicity prevented the use of Venn diagrams in quantum information theory, the introduction of negative entropies has reinstated this useful tool [20,21,15,16]. In particular, we can see how

$$S(L) \not\leq S(LR) . \quad (5)$$

is possible in Fig. 1 if $S(L|R)$ is negative.

The repercussions of such an information-theoretic description of entanglement for the extraction of information from such a state (a measurement) are manifold. Here, we focus on EPR experiments and other quantum paradoxes, and on implications for physical as well as quantum pictures of reality.

3 Information Theory of EPR Experiments

In order to assess the relation between quantum and physical reality in an EPR measurement, we need to describe both the quantum system (the EPR wave-function) *and* the classical devices it becomes entangled with, using information theory.

Fig. 1. Quantum entropy diagrams. (a) Definition of joint $[S(LR)]$ (the total area), marginal $[S(L)$ or $S(R)]$, conditional $[S(L|R)$ or $S(R|L)]$ and mutual $[S(L : R)]$ entropies for a quantum system LR separated into two subsystems L and R ; (b) their respective values for the EPR pair.

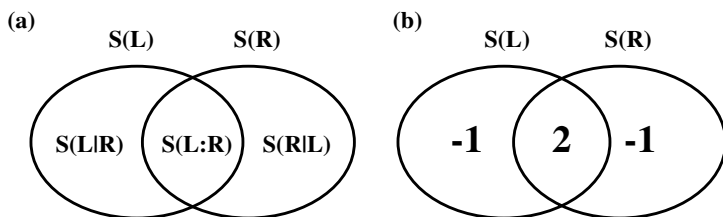


Fig. 2. Measurement of EPR pair Q_1Q_2 by devices A_1 and A_2 .

Let A_1 and A_2 denote measurement *devices*, each of the devices measuring the z component of one member of an EPR pair, for example (see Fig. 2).

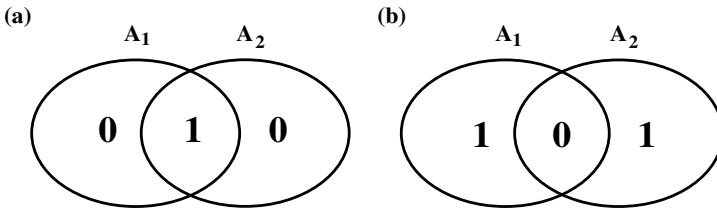
It is an experimental fact that the measurement of the state of one of the particles (say, σ_z) allows a 100% accurate prediction of what the outcome of the measurement of the other one will be. Thus, the outcomes of the measurement of σ_z are perfectly correlated, a situation described by the entropy diagram in Fig. 3a, which appears perfectly classical (no negative numbers appear).

Note that the correlations between the devices are quite unlike those of the quantum system that is measured, a peculiarity that is quantitatively manifested when comparing Figs. 1b and 3a. The reason why the correlations between the measurement devices (Fig. 3a) *incompletely* mirror the entanglement present in the quantum state (Fig. 1b) must be due to the device's classical nature: classical conditional entropies cannot be negative. However, classicality must not be *assumed* for the devices, it is a mathematical result of the information-theoretic treatment of measurement (which gives rise to Fig. 3a) [16].

Assume now that *orthogonal* spin projections are measured on the two halves of the EPR pair, say σ_z on the left particle, and σ_x on the right one. If we assume that measuring the state of one partner confers reality to the state of the measured *system*, we must conclude that the experiment just described would allow us to infer the z and x projections *simultaneously*, a state of affairs strictly forbidden by the uncertainty relation. In their landmark paper [1] EPR there-

fore conclude that, since conventional quantum mechanics cannot describe this peculiar situation, the theory must necessarily be incomplete. This is the essence of the EPR paradox. It relies on a definition of reality based on “certain prediction”¹ according to which the state of the second particle would acquire physical reality after measuring its EPR partner. In fact, for this particular experiment (measuring σ_z on the left and σ_x on the right particle) it is found that the outcomes recorded by the devices are completely *uncorrelated* as depicted by the classical entropy diagram for the *devices* pictured in Fig. 3b. Rather than reflecting an incompleteness of the formalism, these outcomes are *predicted* by quantum information theory, and imply that physical reality is attributed to the state of the second measurement *device*, or more precisely the *relative* state of the devices, while there *cannot* be any correlation between the apparatus and the quantum state proper (as we show below). In view of the importance of this conclusion, let us repeat it once more. Quantum information theory predicts that in EPR-type measurements, the measurement device *cannot* reflect the state of the quantum system. In the language of Mermin [17], the correlations between the devices are real, i.e., possess physical reality, while the quantum system itself does not.

Fig. 3. Entropy diagram for the *devices*: (a) recording σ_z for each of the entangled particles, or (b) recording σ_z for one and σ_x for the other particle.



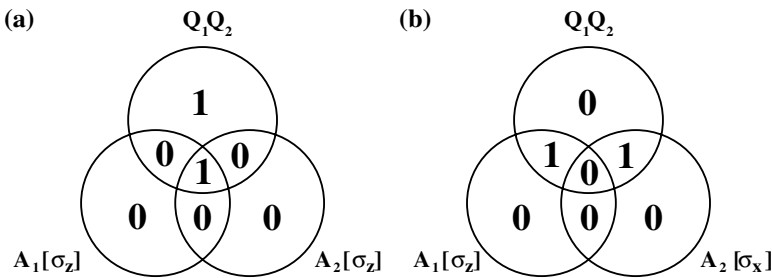
Let us show this in more detail. For a proper quantum information-theoretic analysis, we need to consider four systems: a quantum pair Q_1Q_2 and a pair of ancillae A_1A_2 . The ancillae can be thought of as classical devices that are built to reflect the quantum states. From a measurement point of view, we are interested in the correlations between the *ancillae*, as only such correlations are experimentally accessible (relative states). Before we analyze them using *quantum* entropy diagrams, let us ponder what we expect to find from an orthodox point of view.

One of the fundamental tenets of classical measurement theory is that a measurement device is constructed such as to mirror the state of the object to be measured as accurately as possible. In other words, measurement entails

¹ EPR wrote in [1]: “If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.”

the transfer of this information to a macroscopic system that is more suited to accurate observation, without altering the state of the system. While it is well-known that *quantum* measurements cannot be made without altering the quantum state [22], the general belief is that the quantum state *after* measurement is truthfully portrayed by the device. In other words, it is believed that correlations between the quantum state and the ancillae in the measurement situation allow the extraction of information about the quantum system. Let us consider the “orthodox” (classical) entropy diagram (Fig. 4) for an EPR measurement, drawing the quantum system Q_1Q_2 as one system, measured by the ancillae A_1 and A_2 . These diagrams reveal the paradox inherent in this description. On the

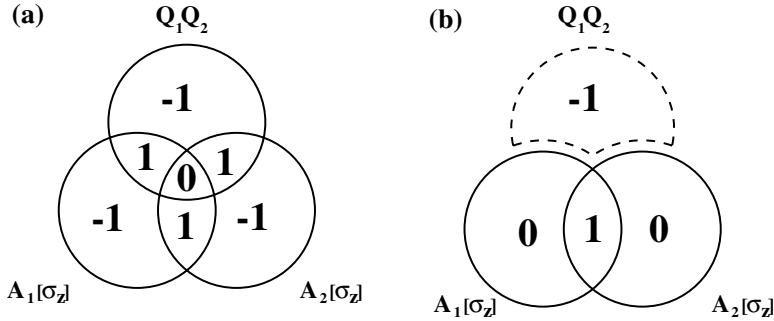
Fig. 4. *Classical* entropy diagram for the EPR measurement of spin-projections: (a) both devices measure σ_z , (b) one device measures σ_z , the other σ_x . Note that the entropy of A_1 and A_2 have to be one bit in each case, as the measurement outcomes are equiprobable, while Q_1Q_2 is thought to have *two* independent equiprobable degrees of freedom.



one hand, when the same projection of the spin (e.g., σ_z) is measured for *both* particles (Fig. 4a) classical reasoning suggests that the quantum system and the measurement devices *share* information (one bit in the center of the diagram). On the other hand, when orthogonal polarizations are measured (Fig. 4b) the measurement devices must appear uncorrelated. According to a “physical realism” or “hidden variable” picture, both diagrams in Fig. 4 must have a common underlying classical diagram relating five ensembles: the EPR pair Q_1Q_2 and the four possible measurements $A_1[\sigma_z]$, $A_1[\sigma_x]$, $A_2[\sigma_z]$, and $A_2[\sigma_x]$ ². This underlying diagram, however, is in contradiction with the Heisenberg uncertainty principle, as it implies that the *counterfactual* variables σ_x and σ_z (of the same particle) can be measured together. Thus, this classical treatment of information leads to a paradox.

² The diagrams in Fig. 4 are obtained from such an underlying diagram by ignoring two out of the five variables: Fig. 4a by ignoring $A_1[\sigma_x]$ and $A_2[\sigma_x]$, Fig. 4b by ignoring $A_1[\sigma_z]$ and $A_2[\sigma_z]$. “Ignoring” a system is achieved by the mathematical operation of tracing it out of the joint density matrix.

Fig. 5. (a) *Quantum* entropy diagram for the EPR measurement of same spin-projections: e.g., A_1 and A_2 both measure σ_z . (b) reduced diagram obtained by tracing over the quantum states Q_1 and Q_2 .



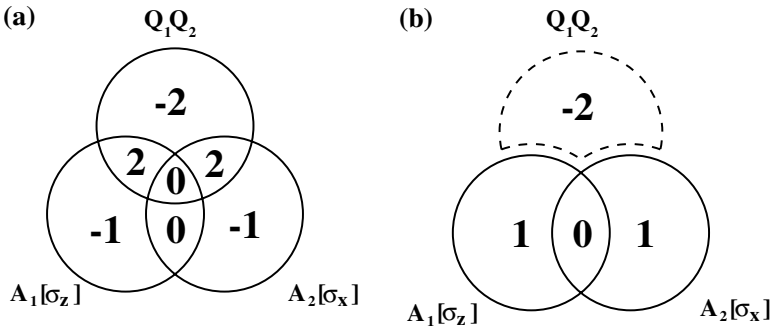
The paradox is resolved by drawing the *quantum* entropy diagrams for the measurements (Figs. 5 and 6). The values for the respective quantum entropies entering these diagrams can be obtained by straightforward calculation [16]. In Fig. 5 the entropy diagram representing the situation where the same polarizations are measured is that of a GHZ state [23]: fully symmetric and maximum quantum entanglement between three entities. As is well-known, tracing over (ignoring) one member of such a triplet produces classical correlations (of the type depicted in Fig. 3a) in the remaining doublet, as indicated in Fig. 5b. As a consequence, the quantum entropy diagram of Fig. 5a *correctly* reproduces the observed correlations between the detectors A_1 and A_2 . Closer inspection of Fig. 5a, however, reveals that while the measurement devices are perfectly correlated as they should, their mutual entropy (the single bit of information gained in the measurement) is *not* shared by the quantum system Q_1Q_2 . In Figs. 5 and 6 this ternary mutual information³ is represented by the center of the diagram, and measures how much of the correlations between the measurement devices is shared by the quantum system. If the center of the diagram is zero, we must conclude that no information is shared between quantum system and classical devices.

The same is true for the measurement situation in Fig. 6a, where *incompatible* polarizations are measured. Again, the (four part) system is fully entangled, and ignoring the quantum state produces the experimentally observed independence of the measurement results (Fig. 6b, compare Fig. 3b). Yet, the correlation between quantum state and measurement device (the mutual information between

³ Just like any entropy, *information*, which is the mutual entropy between *two* systems, can be split up into a conditional and a mutual piece with respect to a third system [13].

the measuring and the measured system) is *unchanged* from the previous arrangement, in fact, it vanishes in both cases⁴.

Fig. 6. (a) Quantum entropy diagram for the EPR measurement of orthogonal spin-projections, e.g., A_1 measures σ_x while A_2 records σ_z . (b) Reduced diagram as above.



This situation leads us to suggest that we must abandon at least one cherished notion of orthodox measurement theory: that the apparatus necessarily reflects the state of the system it was built to measure, by being *correlated* with it in the sense of Shannon. Rather, it is the correlations *between* the ancillae (the reality of their *relative* state) that create the *illusion* of measurement. Indeed, any subsequent measurement on each side (left or right), for example, would yield the *same* result, over and over again, while still not implying *anything* about the quantum wavefunction. Each observer that repeats a measurement becomes classically correlated to the earlier outcome, *whatever the outcome*. Still, the quantum reality of the superposition is unperturbed by these measurements: none of the repeatable measurements yield any information about the quantum state, while they are internally completely consistent. Note that the orthodox interpretation of these correlations involves the concept of a wavefunction collapse: the measurement of the first particle projects—or collapses—the wavefunction of the other one, to account for the perfect correlation. Since the devices do not reflect the state of the quantum system, however, no collapse is needed to explain the correlations, nor does it actually occur, as we now show.

4 Information Theory of Schrödinger Cats

The Schrödinger-cat paradox is of prime importance for the understanding of quantum decoherence and the quantum-classical boundary. The latter have received increased attention recently due to their importance for the design of quantum computation and communication devices [24].

⁴ The mutual information between quantum system and both classical devices also vanishes for any intermediate situation between Figs. 5 and 6, since the joint system $Q_1Q_2A_1A_2$ is always a pure state [16].

The Schrödinger-cat paradox explores the relationship between classical and quantum variables by coupling them together in such a way that the decay of a radioactive substance (say, one isolated atom) implies the demise of a cat locked up with the deadly contraption in a sealed room. The quantum reality of the (isolated) atomic system is that of a superposition of a decayed atom with gamma ray, and an undecayed atom without. If brought into contact with the cat, however, quantum mechanics forces us to include the cat in this entangled wavefunction

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|A^*, 0, L\rangle + |A, 1, D\rangle) , \quad (6)$$

where $|A^*\rangle$ and $|0\rangle$ refer to the excited atom and *absent* gamma, while $|A\rangle$ and $|1\rangle$ are the wavefunctions of the decayed atom and the gamma. Furthermore, $|L\rangle$ and $|D\rangle$ refer to the “live” and “dead” cat eigenstates. The paradox arises if an observer peeks into the room to observe the state of the cat. Does the cat’s wavefunction immediately collapse into one of its eigenstates (dead or alive) upon observation? The preceding analysis teaches us that this is not necessary. The observer can be thought of as a fourth system that is now *quantum entangled* with the previous troika: atom, gamma, and cat

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|A^*, 0, L, l\rangle + |A, 1, D, d\rangle) , \quad (7)$$

introducing “observer eigenstates” $|l\rangle$ and $|d\rangle$. Then, upon tracing over the quantum degrees of freedom of the atom (after all, this experiment involves monitoring the cat and not the atom), the cat (serving as the hapless gamma-ray detector) appears perfectly correlated with the observer peeking into the room. Cat and observer agree, so to speak, about the observation, and their state is entirely classical. Yet, their agreement is completely decorrelated, *disjoint*, from the quantum state, as their *mutual* information shared with the atomic system vanishes. In other words, the classical reality displayed by cat and observer does not imply anything about the quantum reality of atom and gamma ray, or vice versa. Fundamentally, the reason why the observer does not register a cat mired in a quantum superposition of the living and non-living states is because the observer, having interacted with the cat, is entangled with, and thus part of, the *same* wavefunction. As the wavefunction is *indivisible*, an observer (or measurement device) would have to monitor *itself* in order to learn about the wavefunction. This is logically impossible.

5 Conclusions

To summarize, we assert that quantum reality is “real” in the sense that quantum mechanics completely and deterministically describes the evolution of a closed system (not just its wavefunction), and that the statistical character arises from the fact that an observer, because he is part of the closed system, is offered an *incomplete* view of the quantum system he attempts to measure. Consequently, the quantum universe is deterministic as Einstein’s physical reality demands,

but must include the observer as one of its parts due to the inseparability of entangled quantum states. The recent information-theoretic analysis of quantum measurement [15,16] shows that such an observer indeed perceives the system he is measuring as probabilistic, and thus that Bohr's complementarity principle emphasizing the importance of the system/observer relation therefore holds at the same time. If quantum reality is so elusive, how then can we learn about its nature? Fortunately, while negative entropy cannot be reflected in classical instruments directly, it is possible to infer it from combined measurements and comparison with classical bounds (a case in point are Bell inequalities [9], see also [23]). Thus, quantum reality does leave its traces in experiments, while the direct observation of superpositions is impossible.

Acknowledgments

This work was supported in part by NSF Grant Nos. PHY 94-12818 and PHY 94-20470 and by a grant from DARPA/ARO through the QUIC program (#DAAH 04-96-1-3086). N.J.C. is Collaborateur Scientifique of the Belgian National Fund for Scientific Research. An earlier version of this paper was circulated in the Fall of 1996 under the title "Physical Reality and Quantum Paradoxes".

References

1. A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777 (1935).
2. N. Bohr, The quantum postulate and the recent development of atomic theory, *Nature* **121**, 580 (1928).
3. J.A. Wheeler and W.H. Zurek, eds., *Quantum Theory and Measurement* (Princeton University Press, 1983).
4. J.S. Bell, *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987).
5. W. Schommers, ed., *Quantum Theory and Pictures of Reality*, (Springer, Berlin, 1989).
6. P. Busch, P.J. Lahti, and P. Mittelstädt, *The Quantum Theory of Measurement* (Springer, New York, 1991).
7. J.T. Cushing, *Quantum Mechanics—Historical Contingencies and the Copenhagen Hegemony*, (University of Chicago Press, Chicago, 1994).
8. D. Bohm, A suggested interpretation of the quantum theory in terms of hidden variables, I and II, *Phys. Rev.* **85**, 166 (1952).
9. J.S. Bell, On the Einstein Podolsky Rosen paradox, *Physics* **1**, 195 (1965).
10. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer Verlag, Berlin, 1932).
11. E. Schrödinger, Die gegenwärtige Situation in der Quantenmechanik, *Naturwissenschaften* **23**, 807 (1935).
12. C. Monroe, D. M. Meekhof, B.E. King, and D. J. Wineland, A Schrödinger cat superposition state of an atom, *Science* **272**, 1131 (1996); J. J. Slosser and G.J. Milburn, Creating metastable Schrödinger cat states, *Phys. Rev. Lett.* **75**, 418 (1995).

13. C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, 1949).
14. N. Bohr, Can quantum-mechanical description of physical reality be considered complete?, *Phys. Rev.* **48**, 696 (1935).
15. N.J. Cerf and C. Adami, Quantum mechanics of measurement, eprint quant-ph/9605002, unpublished.
16. N.J. Cerf and C. Adami, Information theory of quantum entanglement and measurement, *Physica D* (1998).
17. N.D. Mermin, What is quantum mechanics trying to tell us?, eprint quant-ph/9801057.
18. D. Bohm, *Quantum Theory*, (Prentice-Hall, Englewood Cliffs, 1951), pp. 611-623.
19. A. Wehrl, General properties of entropy, *Rev. Mod. Phys.* **50**, 221 (1978).
20. N.J. Cerf and C. Adami, Negative entropy and information in quantum mechanics, *Phys. Rev. Lett.* **79** (1997) 5194.
21. N.J. Cerf and C. Adami, Negative entropy in quantum information theory, in *New Developments on Fundamental Problems in Quantum Physics*, Fundamental Theories of Physics **81**, M. Ferrero and A. van der Merwe, eds. (Kluwer Academic Publishers, Dordrecht, 1997) p. 77.
22. This is the essence of the quantum non-cloning theorem, see W.K. Wootters and W.H. Zurek, A single quantum cannot be cloned, *Nature* **299**, 802 (1982); D. Dieks, Communication by EPR devices, *Phys. Lett.* **92A**, 271 (1982).
23. D.M. Greenberger, M.A. Horne, and A. Zeilinger, Going beyond Bell's theorem, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, M. Kafatos, ed., (Kluwer, Dordrecht, 1989) p. 69; N.D. Mermin, Quantum mysteries revisited, *Am. J. Phys.* **58**, 731 (1990).
24. D.P. DiVincenzo, Quantum computation, *Science* **270**, 255 (1995); I.L. Chuang, R. Laflamme, P.W. Shor, and W.H. Zurek, Quantum computers, factoring, and decoherence, *ibid.*, p. 1633; C.H. Bennett, Quantum information and computation, *Phys. Today* **48**, 24 (October, 1995).
25. N.J. Cerf and C. Adami, Entropic Bell inequalities, *Phys. Rev. A* **55**, 3371 (1997).

Quantum Generalization of Conditional Entropy and Information

L.B. Levitin

ECE Department
8 St. Mary's St.
Boston University, Boston, MA, 02215
levitin@engr.bu.edu

Abstract. The concepts of conditional entropy of a physical system given the state of another system and of information in a physical system about another one are generalized for quantum one is that the entropy and information in quantum systems. The fundamental difference between the classical case and the quantum one is that the entropy and information in quantum systems depend on the choice of measurements performed over the systems. It is shown that some equalities of the classical information theory turn into inequalities for the generalized quantities. Specific quantum phenomena such as EPR pairs and "superdense coding" are described and explained in terms of the generalized conditional entropy and information.

Keywords: Quantum Information, Quantum Conditional Entropy, Information in Entangled States, Quantum Measurement

The purpose of this paper is to show how the concepts of conditional entropy and information (by C.E. Shannon) can be generalized for quantum systems. It is shown that the freedom of choice of measurement performed over quantum systems results in some classical information-theoretical equalities becoming inequalities. The generalized concepts are applied to describe typical quantum-mechanical phenomena of EPR pairs and the so called "superdense coding".

Consider a quantum system that consist of two subsystems, A and B . Let α and β be complete sets of variables in A and B , respectively. Let $\{a\}$ and $\{b\}$ be sets of values taken on by α and β , respectively. The Hilbert space of the states of the system \mathcal{H} is the tensor product of the Hilbert spaces of the subsystems: $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. The state of the system is, in general, a mixed one and is described by a joint density matrix in \mathcal{H} :

$$\rho(A, B) = \|\rho_{ab, a'b'}\| \quad (1)$$

where the matrix elements are written in the basis formed by tensor products of eigenvectors of α and β . The marginal density matrices of the subsystems are given by partial traces:

$$\begin{aligned}\rho(A) &= \text{Tr}_B \rho(A, B) = \|\sum_b \rho_{ab, a'b}\| \\ \rho(B) &= \text{Tr}_A \rho(A, B) = \|\sum_a \rho_{ab, ab'}\|\end{aligned}\quad (2)$$

The joint entropy of the system and the marginal entropies of the subsystems are, respectively,

$$\begin{aligned}H(A, B) &= -\text{Tr} \rho(A, B) \log \rho(A, B) \\ H(A) &= -\text{Tr} \rho(A) \log \rho(A) \\ H(B) &= -\text{Tr} \rho(B) \log \rho(B)\end{aligned}\quad (3)$$

Denote by $H(\alpha, \beta)$, $H(\alpha)$, and $H(\beta)$ the entropies of the results of the measurements performed over the states $\rho(A, B)$, $\rho(A)$ and $\rho(B)$, respectively, in the bases formed by eigenvectors of α and β . Then, according to Klein's Lemma[1,2],

$$H(A, B) \leq H(\alpha, \beta); \quad H(A) \leq H(\alpha); \quad H(B) \leq H(\beta) \quad (4)$$

where the inequalities hold iff the corresponding density matrix is diagonal in the measurement basis. Inequalities (4) express, by von Neuman[2], the effect of the irreversibility of quantum measurements.

The next step is to introduce conditional density matrix [3,4]. Denote by P_b the projection operator on the eigenvector corresponding to eigenvalue b of the complete set of variables β . Then the **conditional density matrix** of the system A , given that the result of measurement of β performed over system B is b , is

$$\rho(A|\beta = b) = \frac{P_b \rho(A, B) P_b}{\text{Tr}_A \rho(A, B) P_b} = \left\| \frac{\rho_{ab, a'b}}{\sum_a \rho_{ab, ab}} \right\| \quad (5)$$

Note that the denominator in (5) is just the probability of β to take on value b ,

$$\text{Tr}_A \rho(A, B) P_b = \text{Pr}\{\beta = b\} \quad (6)$$

Also, for any b , $\text{Tr}_A \rho(A|\beta = b) = 1$ (as it should be for a density matrix), and

$$\sum_b (\text{Tr}_A \rho(A, B) P_b) \rho(A, |\beta = b) = \rho(A) \quad (7)$$

Then, in accordance with Shannon's definition of conditional entropy, and the quantum definition of entropy[2], the conditional entropy of system A given the measurement β performed over B , is

$$H(A|\beta) = - \sum_b (\text{Tr}_A \rho(A, B) P_b) \text{Tr}_A \rho(A|\beta = b) \log \rho(A|\beta = b) \quad (8)$$

By Klein's Lemma, for any α and β ,

$$0 \leq H(A|\beta) \leq H(\alpha|\beta) \quad (9)$$

where the equality holds iff all $\rho(A|\beta = b)$ commute and are diagonal in the basis of eigenvectors of α .

The quantity given by (8) depends on the choice of measurement β (or, more exactly, on the choice of basis in \mathcal{H}_B formed by the eigenvectors of the complete set of variables β). This situation is different from that in classical theory, where any complete set of variables would give the same conditional entropy for the system A . Since the meaning of conditional entropy is to express the uncertainty of the subsystem A under the constraints imposed by any possible measurement performed over the subsystems B , we suggest the following definition.

Def. 1. Conditional entropy of a subsystem A , conditioned by a subsystem B , is

$$H(A|B) = \inf_{\beta} H(A|\beta) \quad (10)$$

Theorem 1.

$$H(A|B) \leq H(A)$$

Proof.

$$H(\alpha|\beta) \leq H(\alpha)$$

$$\inf_{\beta} H(\alpha|\beta) \leq H(\alpha)$$

$$H(A|B) = \inf_{\beta} \sum_b \inf_{\alpha} H(\alpha|\beta = b) \leq \inf_{\alpha} \inf_{\beta} H(\alpha|\beta) \leq \inf_{\alpha} H(\alpha) = H(A) \quad \square$$

Note that according to classical information theory $H(\alpha|\beta) = H(\beta) + H(\alpha|\beta)$. However, this equality turns into inequality for quantum systems.

Theorem 2.

$$H(A, B) \leq H(B) + H(A|B) \quad (11)$$

The inequality (11) stems from the fact that $H(A, B) = \inf_{\gamma} H(\gamma)$, where γ is a basis in \mathcal{H} , which is not necessarily a tensor product of bases in \mathcal{H}_A and \mathcal{H}_B .

Turning to information, we should remember that information (in Shannon's sense) is defined iff there is a pair of random variables. In the case of quantum physics, these random variables are observables, and their values are outcomes of measurements. Thus, if α and β are complete sets of observables for systems A and B , respectively, we may consider mutual information between the outcomes of measurements of α and β :

$$I(\alpha; \beta) = H(\alpha) - H(\alpha|\beta) \quad (12)$$

Def. 2. Information in the subsystem B about the subsystem A (and vice versa) is

$$I(A; B) = \sup_{\alpha, \beta} I(\alpha; \beta) \quad (13)$$

Then the classical equality (12) turns into inequality.

Theorem 3.

$$I(A; B) \leq H(A) - H(A|B) \quad (14)$$

The equality in (14) holds iff for β such that $H(A|B) = H(A|\beta)$ all $\rho(A|\beta = b)$ commute and is achieved for the basis of eigenvectors of α such that all $\rho(A|\beta = b)$ are diagonal there.

Proof. It follows from the entropy defect principle [5,6] that for any β

$$I(A; B) = \sup_{\alpha} I(\alpha; \beta) \leq H(A) - H(A|\beta) \quad (15)$$

where the equality holds iff there exists a basis of eigenvectors of α where all $\rho(A|\beta = b)$ are diagonal.

From (15),

$$I(A; B) = \sup_{\alpha, \beta} I(\alpha; \beta) \leq \sup_{\beta} [H(A) - H(A|\beta)] = H(A) - \inf_{\beta} H(A|\beta) = H(A) - H(A|B) \quad \square$$

Example 1: EPR pair

Consider two particles, A and B , each having a 2-dimensional Hilbert space. Let $\{|0\rangle, |1\rangle\}$ be the basis in \mathcal{H}_A and the basis in \mathcal{H}_B formed by eigenvectors of α and eigenvectors of β , respectively. Let the system be in pure entangled state with a density matrix

$$\rho(A, B) = \frac{1}{2}(|01\rangle - |10\rangle)(\langle 01| - \langle 10|) \quad (16)$$

In other words, the density matrix in the (α, β) basis is

$$\rho(A, B) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The marginal density matrices are :

$$\rho(A) = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad \rho(B) = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (17)$$

The conditional density matrices of the subsystem A are:

$$\rho(A|\beta = 0) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}; \quad \rho(A|\beta = 1) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (18)$$

Hence $H(A, B) = 0$ and $H(\alpha, \beta) = 1$; $H(A) = H(\alpha) = 1$; $H(B) = H(\beta) = 1$; $H(A|B) = H(A|\beta) = 0$.

Note that all $\rho(A|\beta = b)$ commute and are diagonal in basis β . Therefore,

$$I(A; B) = \sup_{\alpha, \beta} I(\alpha; \beta) = H(A) - H(A|B) = 1 \quad (19)$$

Let us now take a different basis in B , formed by eigenvectors of another observable $\tilde{\beta}$:

$$\left\{ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right\} \quad (20)$$

In $(\alpha, \tilde{\beta})$ basis

$$\rho(A, B) = \frac{1}{4} \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix} \quad (21)$$

Then

$$\rho(A|\tilde{\beta} = 0) = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}; \quad \rho(A|\tilde{\beta} = 1) = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (22)$$

$$\text{Hence,} \quad H(\alpha|\tilde{\beta}) = 1 \quad \text{and} \quad I(\alpha; \tilde{\beta}) = 0 \quad (23)$$

But the conditional density matrices (22) still commute, and $H(A|\tilde{\beta}) = H(A|B) = 0$. Therefore there should exist an observable $\tilde{\alpha}$ in A such that

$$I(A; B) = I(\tilde{\alpha}; \tilde{\beta}) = 1 \quad (24)$$

Indeed, this is true for an observable $\tilde{\alpha}$ with eigenvectors

$$\left\{ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right\} \quad (25)$$

Example 2. Superdense Coding

The phenomenon of superdense coding[7] is essentially quantum-mechanical and has no classical counterpart. It can be described as follows.

Suppose we have an EPR pair of particles B and B' in an entangled state (16) and a third particle of the same sort, A , whose initial state is independent of the pair (B, B') . Then, by local unitary transformations over the particles A and B' one can obtain four different entangled states of particles A and B depending on the value of a classical random variable C which takes four equiprobable values 0,1,2,3. Namely

$$\begin{aligned} C = 0 : \psi(A, B) &= \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \\ C = 1 : \psi(A, B) &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \\ C = 2 : \psi(A, B) &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\ C = 3 : \psi(A, B) &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \end{aligned} \quad (26)$$

Note that states(26) are pure and orthogonal and, therefore, form a basis in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ consisting of eigenvectors of an observable γ in \mathcal{H} .

Since C is a classical variable, the joint density matrix $\rho(A, B, C)$ must be diagonal in C . In the basis $(\alpha, \beta) = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ the density matrix has the following form:

$$\rho(A, B, C) = \frac{1}{8} \begin{bmatrix} 0 & 0 & 0 & 0 & & & & \\ 0 & 1 & 1 & 0 & & & & \\ 0 & 1 & 1 & 0 & & & & \\ 0 & 0 & 0 & 0 & & & & \\ & & & & 0 & 0 & 0 & 0 \\ & & & & 0 & 1 & -1 & 0 \\ & & & & 0 & -1 & 1 & 0 \\ & & & & 0 & 0 & 0 & 0 \\ & & & & & & & 1 & 0 & 0 & 1 \\ & & & & & & & 0 & 0 & 0 & 0 \\ & & & & & & & 0 & 0 & 0 & 0 \\ & & & & & & & 1 & 0 & 0 & 1 \\ & & & & & & & & & 1 & 0 & 0 & -1 \\ & & & & & & & & & 0 & 0 & 0 & 0 \\ & & & & & & & & & 0 & 0 & 0 & 0 \\ & & & & & & & & & -1 & 0 & 0 & 1 \end{bmatrix} \quad (27)$$

(The empty blocks, in the above matrix, are all zero.)

Now it is easy to see that subsystems A and B taken separately carry no information about C :

$$I(A; C) = I(B; C) = I(\alpha; C) = I(\beta; C) = 0 \quad (28)$$

The measurement of the observables in both A and B yields only one bit of information:

$$I(\alpha, \beta; C) = 1 \quad (29)$$

as it should be in the case of three random variables in classical information theory. However, the measurement of the variable γ , whose eigenvectors form a basis in the \mathcal{H} which is **not** a tensor product of bases in \mathcal{H}_A and \mathcal{H}_B , yields, paradoxically, two bits of information, thereby identifying uniquely the value of C :

$$I(A, B; C) = I(\gamma; C) = 2 \quad (30)$$

This is, indeed, a paradox, since in classical information theory, if α and β are binary random variables, the joint information in (α, β) about C is upper-bounded as follows:

$$I(\alpha, \beta; C) = I(\alpha; C) + I(\beta; C|\alpha) \leq I(\alpha; C) + H(\beta|\alpha) \quad (31)$$

Taking into account that $H(\beta|\alpha) \leq H(\beta) \leq 1$, and, in our case, $I(\alpha; C) = 0$, we obtain that $I(\alpha, \beta; C) \leq 1$. The example of superdense coding shows that the joint information in two two-dimensional quantum systems A and B about random variable C $I(A, B; C) = 2 > 1$, inspite of the fact that $I(A; C) = I(B; C) = 0$. Since in our case $H(B|A) = 1$, we conclude that an inequality opposite to (31) holds here, namely, that

$$I(A, B; C) > I(A; C) + H(B|A) \quad (32)$$

which resolves the paradox.

These unusual properties of information between quantum systems results from the fact that, in contrast with classical theory, this information depends on the choice of measurement, i.e. on the choice of a complete set of one-dimensional orthogonal projection operators in the Hilbert space of the system. Let us point out in conclusion that definitions of conditional entropy and information for quantum systems can be easily modified in an obvious way to include indirect measurements, i.e. complete sets of one-dimensional projection operators which form a non-orthogonal resolution of identity in the Hilbert space.

References

1. O. Klein, Zeitschr. f. Phys., B.72, 1931, 767.
2. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin, 1932
3. V.V. Mityugov, *Physical Principles of Information Theory*, Sovetskoye Radio, Moscow, 1976 (in Russian)
4. R. Balian, *From Microphysics to Macrophysics. Methods and Applications of Statistical Physics*, Vol. 1, Springer-Verlog, Berlin, Hiedelberg, New York, 1991
5. L.B. Levitin, *On the Quantum Measure of Information*, Proc. of the 4th Conf. on Information Theory, Tashkent, 1969, 111-116. English Translation: Annales de Fondation Louis de Broglie, v.21, No. 3, 1996, 345-348.
6. A.S. Holevo, *Some Estimates of Information Transmitted over Quantum Communication Channel*, Probl. of Inf. Transm., v. 9, No. 3, 1973, 3-11
7. C.H. Bennett and S.J. Wiesner, Phys. Rev. Letters, v. 69, 1992, 2881

Accessible Information in Multi-access Quantum Channels

A.E. Allahverdyan and D.B. Saakian

Yerevan Physics Institute
Alikhanian Brothers St.2, Yerevan 375036, Armenia

Abstract. The accessible information in multi-access quantum channels are considered. Classical messages from independent sources, which are represented as some quantum states, are transported by a channel to one address. The messages can interact with each other and with external environment. After statement of problem and proving some general results we investigate physically important case when information is transported by states of an electromagnetic field. One-way communication by noisy quantum channels is also considered.

1 Introduction

Physical ideas played important role as sources for information theory [19]. Investigation of quantum mechanical aspects of information theory was started by Gordon, Lebedev and Levitin [2], and other researchers (see references in [1]). Now it is well known that quantum-mechanical and thermodynamical limitations are very important for correct consideration of information-theoretical models [19].

More generally, quantum information theory contains two distinct types of problems. The first type describes transmission of classical information through a quantum channel (the channel can be noisy or noiseless). In such scheme bits encoded as some quantum states and only this states or its tensor products are transmitted. In the second case arbitrary superposition of this states or entanglement states can be transmitted. In the first case the problems can be solved by methods of classical information theory, but in the second case new physical representations are needed. In this work we investigate the problems of the first type.

In the almost all paper devoted to problems of physical information theory only one-way communication are considered- i.e., there is one input with some initial quantum ensemble and one output where quantum states are detected. Here quantum ensemble is some set of quantum states (which can be represented by corresponding density matrices) with corresponding probabilities. But in the practice multi-terminal communication schemes also are important. In this case there are several input for information and several output for detection. In this communication scheme messages are represented as some physical systems and

there are interactions between the messages and external environment. For general discussion about multi-terminal classical, mathematical information theory see [10].

The paper is organized as follows. In section 2 we discuss the general statement of the problem, and derive some upper bounds for the accessible information like Holevo bound in one-way communication [4]. In section 3 practically important case is considered when information is transmitted by coherent and quadrature-squeezed states of an electromagnetic field.

2 The Upper Bounds for Accessible Information.

We consider one output which receive information from two independent sources. The messages of these sources are represented as some quantum states. More exactly we can say that for any letter of the classical alphabets the concrete quantum state is generated. The initial quantum ensembles of two independent sources are

$$\rho^{(1)} = \sum_{\alpha} p_{\alpha}^{(1)} \rho_{\alpha}^{(1)}, \quad (1)$$

$$\rho^{(2)} = \sum_{\beta} p_{\beta}^{(2)} \rho_{\beta}^{(2)}. \quad (2)$$

After initial preparation the quantum states of (1, 2) penetrate through a quantum channel. In this channel there are interactions between states of the sources and an interaction with the environment. The possible concrete mechanisms of these interactions will be discussed later. The general effect of the noisy quantum channel can be described by a quantum evolution operator \hat{S} with kraussian representation

$$\hat{S}\rho = \sum_{\mu} A_{\mu}^{\dagger} \rho A_{\mu}, \quad \sum_{\mu} A_{\mu} A_{\mu}^{\dagger} = \hat{1}. \quad (3)$$

This operators must be linear, completely positive and trace-preserving [11, 7]. After interaction, the receiver has the states $\sigma_{\alpha\beta}^{(12)}$

$$\hat{S}(\rho_{\alpha}^{(1)} \otimes \rho_{\beta}^{(2)}) = \sigma_{\alpha\beta}^{(12)} \quad (4)$$

These states are contained in the quantum ensemble

$$\sigma = \sum_{\alpha, \beta} p_{\alpha} p_{\beta} \sigma_{\alpha\beta}^{(12)} \quad (5)$$

At the output of the channel the receiver should separate and recognize the messages of each source. Besides a noise which is introduced by the environment each transmitter also acts as a noise for other. Now the receiver needs some measurement procedure. It is important that the elements of (1, 2) can be nonorthogonal or nonorthogonality can occur after action of (4). In this case for

more optimal distinguishing between different quantum states we need generalized measurement procedure [5]. This type of measurement is represented by some nonorthogonal resolution of identity

$$\sum_{\gamma} E_{\gamma} = 1, \quad E_{\gamma} > 0 \quad (6)$$

If system with density matrix ρ is measured then probability of the result γ is $\text{tr}(\rho E_{\gamma})$. In [3] was shown that for distinguishing some nonorthogonal states measurements like (6) more optimal than usual. At the output of the channel (4) as result of some generalized measurement like (6) we have the conditional probabilities

$$p(\gamma/\alpha\beta) = \text{tr}(E_{\gamma}\sigma_{\alpha\beta}^{(12)}) \quad (7)$$

This procedure is called decoding. With initial and independent distributions

$$p(\alpha), \quad p(\beta) \quad (8)$$

the dual multi-access channel in the classical sense is determined [10,9].

Let R_1, R_2 are the maximal quantities of information which the sources can transport in the regime of the reliable connection (it is the connection with small probability of an error in the decoding). As was shown in [10] such R_1, R_2 must satisfied the following conditions

$$R_1 < I(\alpha : \gamma/\beta), \quad R_2 < I(\beta : \gamma/\alpha), \quad (9)$$

$$R_1 + R_2 < I(\alpha \otimes \beta : \gamma). \quad (10)$$

Where

$$I(\alpha : \gamma/\beta) = \sum_{\alpha, \beta, \gamma} p(\alpha, \beta, \gamma) \ln \frac{p(\gamma/\alpha\beta)}{p(\gamma/\beta)}, \quad (11)$$

$$I(\alpha \otimes \beta : \gamma) = \sum_{\alpha, \beta, \gamma} p(\alpha, \beta, \gamma) \ln \frac{p(\gamma/\alpha\beta)}{p(\gamma)}. \quad (12)$$

The second value is usual mutual information between ensembles γ and $\alpha \otimes \beta$. The first value is called mutual-conditional information (mc-information). The mutual information of two ensembles is reduction of entropy of one ensemble if the other is observed. Mc-information has the same meaning but after realization of the conditional ensemble. The physical meaning of (9, 10) is follows. Eq. (10) is usual Shannon formula for joint channel, i.e. when considering the pair of symbols presented at both inputs as a single symbol of a larger alphabet. Eq. (9) arises because messages from the different inputs should be separated at the output. Thus the amount of information which can be transported by the first transmitter is limited by mutual information between the first input and the output when the second input is fixed, and analogously for the second transmitter. Now for long sequences in the inputs and memoryless external noise we have mc-information.

The problem of physical information theory in this case is investigation the results of (9), (10) for physically important noisy channels.

In general case the values (11), (12) can not be calculated explicitly. Therefore investigation the upper bounds for this values is an important problem. For the one-way communication such theorems was proved by A.Holevo [4]. The most general results in this direction was obtained in the [6]. We shall obtain the measurement independent upper bounds for (11), (12) by using the methods of [6]. For this purpose we need some general results from quantum statistical physics.

Quantum relative entropy between two density matrices ρ_1, ρ_2 is defined as follows

$$S(\rho_1||\rho_2) = \text{tr}(\rho_1 \log \rho_1 - \rho_1 \log \rho_2). \quad (13)$$

This positive quantity was introduced by Umegaki [14] and characterizes a degree of 'closeness' between density matrices ρ_1, ρ_2 . The properties of quantum relative entropy were reviewed by M.Ohya [13]. Here we need one basic theorem which was proved by Lindblad [12].

$$S(\rho_1||\rho_2) \geq S(\hat{S}\rho_1||\hat{S}\rho_2). \quad (14)$$

I.e. after action of \hat{S} the quantum states can be only more 'close'. For any generalized measurement $\{E_\gamma\}$ and density matrix ρ we define the following transformation

$$\rho \mapsto \text{diag}\{\text{tr}(E_1\rho), \text{tr}(E_2\rho), \dots\} \quad (15)$$

In other words ρ is transformed to a diagonal matrix with the corresponding diagonal elements. This map is also general quantum evolution operator like (3). Now we shall use map (15) and theorem (14) for values like

$$\text{tr}[\hat{S}(\rho_\alpha^{(1)} \otimes \rho_\beta^{(2)}) (\ln \hat{S}(\rho_\alpha^{(1)} \otimes \rho_\beta^{(2)}) - \ln \hat{S}(\rho^{(1)} \otimes \rho^{(2)}))] \quad (16)$$

We get

$$I(\alpha : \gamma/\beta) \leq - \sum_{\alpha, \beta} p_\alpha^{(1)} p_\beta^{(2)} S(\hat{S}(\rho_\alpha^{(1)} \otimes \rho_\beta^{(2)})) + \sum_{\beta} p_\beta^{(2)} S(\hat{S}(\rho^{(1)} \otimes \rho_\beta^{(2)})) \quad (17)$$

$$I(\alpha \otimes \beta : \gamma) \leq - \sum_{\alpha, \beta} p_\alpha^{(1)} p_\beta^{(2)} S(\hat{S}(\rho_\alpha^{(1)} \otimes \rho_\beta^{(2)})) + S(\hat{S}(\rho^{(1)} \otimes \rho^{(2)})) \quad (18)$$

These inequalities are very useful if general limits are developed. Eqs. (17, 18) are the central results of this section. The attainability of this bounds will be discussed elsewhere.

3 Information Transmission by States of an Electromagnetic Field.

In practice the most important tool for information transmission is the electromagnetic field. We briefly recall the connection between the formalism which

is described above and the characteristics of the electromagnetic field in the vacuum or linear dielectric media [1]. Here the information is connected with longitudinal characteristics of a plane wave with fixed center-frequency and narrow bandwidth. But transverse (polarization, wave vector) characteristics are fixed. This statement of question is more or less realizable in the practice.

Two modes of the electromagnetic field with frequencies ω_1, ω_2 penetrate into the noisy channel. We have considered two kinds of communication channels, which are coherent channel and quadrature-squeezed (briefly squeezed) channel. In the first case the inputs of the sources are coherent states of chosen field modes, and information is carried in the pattern of complex amplitude excitations. In the second case the inputs are squeezed states, and information is carried in the pattern of excitations of squeezed quadratures. Relative to a coherent states a quadrature-squeezed states have reduced quantum uncertainty in one quadrature component, called the squeezed quadrature. There is a corresponding increase in the uncertainty in the orthogonal quadrature component, called the amplified quadrature. For the case of coherent states information can be recovered by heterodyne detection, i.e., by measuring both quadratures of the mode. For the case of quadrature-squeezed states information is recovered by measuring of squeezed quadrature. In the channel the modes can interact together and with external thermostat. We assume that the interaction between modes and the interaction with the thermostat are linear (for discussion about realization of this type of interaction see [15]).

We describe our model by quantum Langevin equation [16] where rotating wave approximation is done.

$$i\dot{a}_1 = \omega_1 a_1 - i\frac{\gamma a_1}{2} + k a_2 + i\bar{F}_1, \quad (19)$$

$$i\dot{a}_2 = \omega_2 a_2 - i\frac{\gamma a_2}{2} + k a_1 + i\bar{F}_2 \quad (20)$$

Where a_1, a_2 are annihilation operators for the modes, k is the strength of the cross mode interaction, γ is the damping constant (for simplicity we choose damping constant same for the modes), $\bar{F}_1(t), \bar{F}_2(t)$ are langevin forces (white noise). These equations are generated by Hamiltonian

$$H = \omega_1 a_1^\dagger a_1 + \omega_2 a_2^\dagger a_2 + k(a_1^\dagger a_2 + a_2^\dagger a_1) \quad (21)$$

Solution of these equations can be obtained immediately. For example

$$a_1 = a_1(0)(\epsilon e^{i\phi_1 t} + (1-\epsilon)e^{i\phi_2 t}) - a_2(0)\sqrt{\epsilon(1-\epsilon)}(e^{i\phi_1 t} - e^{i\phi_2 t}) + \sqrt{\epsilon}e^{i\phi_1 t} \int_0^t e^{i\phi_1 t'} F_1(t')dt' + \sqrt{1-\epsilon}e^{i\phi_2 t} \int_0^t e^{i\phi_2 t'} F_2(t')dt' \quad (22)$$

Where:

$$\phi_{1,2} = -\lambda_{1,2} + i\frac{\gamma}{2}, \quad \lambda_{1,2} = \frac{\omega_1 + \omega_2 \pm \sqrt{(\omega_1 - \omega_2)^2 + 4k^2}}{2} \quad (23)$$

$$\frac{1 - \epsilon}{\epsilon} = \frac{(\lambda_1 - \omega_1)^2}{k^2} \quad (24)$$

$$\langle F_i^\dagger(t) F_i(t') \rangle = \gamma \bar{n}_T (\lambda_{1,2}) \delta_{ij} \delta(t - t'), \quad \bar{n}_T = (\exp(\lambda_{1,2}/T) - 1)^{-1} \quad (25)$$

$$\langle F_i(t) F_j^\dagger(t') \rangle = \gamma (\bar{n}_T (\lambda_{1,2}) + 1) \delta_{ij} \delta(t - t'), \quad \langle F_i F_j \rangle = 0 \quad (26)$$

Now we assume that modes a_1, a_2 at the $t = 0$ (in the input of the channel) are in squeezed states with squeezing parameters r_1, r_2 (some of these parameters can be zero) and shifts $(\alpha_1, \alpha_2), (\beta_1, \beta_2)$ (for definition of squeezed states see [11]). In the output of the channel (at the moment t) one of the modes (for example a_1) is measured. The case when two modes are measured, and decoders can interchange the available information can be obtained by simple modification of the final results. We shall consider only two type of measurements. In the case of heterodyning both component of the mode is measured simultaneously: [1]

$$p(\gamma/\alpha\beta) = \frac{1}{\pi} \langle \gamma | \rho(t; a_1) | \gamma \rangle \quad (27)$$

Where $|\gamma\rangle$ is coherent state with shift γ , and $\rho(t; a_1)$ is the density matrix of the mode a_1 at the moment t . In the second case the squeezed component of the mode is measured. In this case

$$p(\gamma_1/\alpha\beta) = \langle \gamma_1 | \rho(t; a_1) | \gamma_1 \rangle \quad (28)$$

Where $|\gamma_1\rangle$ is an eigenstate of the measuring component. Later we shall treat homodyne and heterodyne measurements simultaneously, and shall use symbol γ for the both cases.

Simple analysis shows [10] that capacities are maximized (the maximization is done by all distributions with fixed initial parameters) by gaussian input probabilities

$$p(\alpha) \sim \exp\left(-\frac{1}{2}\alpha^T K_\alpha \alpha\right), \quad p(\beta) \sim \exp\left(-\frac{1}{2}\beta^T K_\beta \beta\right) \quad (29)$$

Where

$$\alpha^T = (\alpha_1, \alpha_2), \quad \beta^T = (\beta_1, \beta_2).$$

Here and in future gaussian integrals will be written up to multiplicative factor. The Langevin equations ([19,20]) are linear and after time t a gaussian state remains the gaussian. It is convenient to calculate ([28,27]) in the formalism of Wigner functions. The Wigner function of general gaussian state can be represented in the following form as function of the means and square-means of this state

$$W(\text{Re}\gamma, \text{Im}\gamma) \sim \exp\left(-\frac{(\text{Re}\gamma - \langle \text{Re}a \rangle)^2}{2} K_{11} - \frac{(\text{Im}\gamma - \langle \text{Im}a \rangle)^2}{2} K_{22} - (\text{Re}\gamma - \langle \text{Re}a \rangle)(\text{Im}\gamma - \langle \text{Im}a \rangle) K_{12}\right) \quad (30)$$

Where

$$\begin{pmatrix} K_{11} & K_{12} \\ K_{12} & K_{22} \end{pmatrix}^{-1} = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix} \quad (31)$$

$$\begin{aligned}
 c_{11} &= \langle (\text{Re}a)^2 \rangle - \langle \text{Re}a \rangle^2 \\
 c_{22} &= \langle (\text{Im}a)^2 \rangle - \langle \text{Im}a \rangle^2 \\
 c_{12} = c_{21} &= \frac{\langle \text{Re}a \text{Im}a + \text{Im}a \text{Re}a \rangle}{2} - \langle \text{Im}a \rangle \langle \text{Re}a \rangle
 \end{aligned} \tag{32}$$

As we see this channel is gaussian. Capacity region for two-terminal gaussian channel can be obtained exactly. If vectors α, β with distribution (29) are initial messages for the channel then we have for output vector γ

$$\gamma = A_1 \alpha + A_2 \beta + z. \tag{33}$$

Where A_1, A_2 are some matrices (we shall use the specific form of these matrices for homodyne and heterodyne measurements), and z is additive gaussian noisy vector with correlation matrix K^{-1}

$$\langle z z^T \rangle = K^{-1} \tag{34}$$

In this case we get

$$\begin{aligned}
 I(\gamma : \alpha \otimes \beta) &= \ln \det(1 + K A_1 K_\alpha^{-1} A_1^T + K A_2 K_\beta^{-1} A_2^T) \\
 I(\gamma : \alpha / \beta) &= \ln \det(1 + K A_1 K_\alpha^{-1} A_1^T) \\
 I(\gamma : \beta / \alpha) &= \ln \det(1 + K A_2 K_\beta^{-1} A_2^T)
 \end{aligned} \tag{35}$$

After some not hard but tedious calculations we come to the concrete results. The first case is heterodyne measurement of the mode 1 at the moment t with

$$r_1 = 0, \quad r_2 = 0. \tag{36}$$

In this case we should choose the following initial distributions

$$K_\alpha^{-1} = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}, \quad K_\beta^{-1} = \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix} \tag{37}$$

With

$$x = \frac{1}{2} \bar{n}_1, \quad y = \frac{1}{2} \bar{n}_2. \tag{38}$$

Where \bar{n}_1, \bar{n}_2 are mean photon number of the initial distributions. In this case we have from general formulas

$$\begin{aligned}
 I(\gamma : \alpha \otimes \beta) &= \ln \left(1 + \frac{\bar{n}_1 e^{-\gamma t} + 2(\bar{n}_2 - \bar{n}_1) \epsilon (1 - \epsilon) e^{-\gamma t} (1 - \cos(t(\lambda_1 - \lambda_2)))}{\frac{1}{2}(e^{-\gamma t} + \Psi + 1)} \right) \\
 I(\gamma : \alpha / \beta) &= \ln \left(1 + \frac{\bar{n}_1 e^{-\gamma t} (1 - 2\epsilon(1 - \epsilon)(1 - \cos(t(\lambda_1 - \lambda_2))))}{\frac{1}{2}(e^{-\gamma t} + \Psi + 1)} \right)
 \end{aligned} \tag{39}$$

And for Ψ we have

$$\Psi = (1 - e^{-\gamma t})(2\epsilon \bar{n}_T(\lambda_1) + 2(1 - \epsilon) \bar{n}_T(\lambda_2) + 1) \tag{40}$$

We don't write expression for $I(\beta : \gamma/\alpha)$ if written formulas are sufficient for understanding the behavior of this quantity.

For $t = 0$ these information measures coincide with well known expression for capacity of coherent state channel. For large t (39) tend to zero. But as we see the decay is not monotonic: the additional oscillation occur due to the interaction between different modes. If we take $\epsilon = 1$ we come to one-terminal channel with gaussian noise. In this case the capacity monotonically tends to zero.

Now about homodyning with arbitrary r . In this case we should choose as initial distributions

$$K_\alpha^{-1} = \begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}, \quad K_\beta^{-1} = \begin{pmatrix} y & 0 \\ 0 & 0 \end{pmatrix} \quad (41)$$

It is product of the gaussian distribution for the measuring component and delta-function for the other component. With the choice of (41) we have optimal distribution of the input energy. In this case we have the following connection between dispersion of the distribution and mean photon number [1]

$$x = \bar{n}_1 - \text{sh}^2 r_1, \quad y = \bar{n}_2 - \text{sh}^2 r_2. \quad (42)$$

In the first we consider one-terminal communication with arbitrary squeezing parameter r . In this case we have

$$I(\alpha : \gamma_1) = \frac{1}{2} \ln \left(1 + \frac{\cos^2(\lambda t)(\bar{n} - \text{sh}^2 r)}{\frac{1}{4}e^{-2r} + 2\sin^2(\lambda t)\text{sh}(2r) + C} \right) \quad (43)$$

For given \bar{n} optimal squeezing parameter is determined by the following formula

$$e^{2r} = \frac{\sqrt{\cos^4(\lambda t) + C^2 + 4C(2\bar{n} + 1)\cos^2(\lambda t)} - \cos^2(\lambda t)}{C} \quad (44)$$

Where

$$C = (e^{\gamma t} - 1)(2\bar{n}_T + 1) \quad (45)$$

As we see optimal time-dependent squeezing parameter tends to zero for large t . It is well known that for zero t the squeezed-state channel are more effective than the coherent-state one [1]. Indeed we have

$$I = \ln(1 + \bar{n}) \quad (46)$$

for the coherent-state channel and

$$I = \ln(1 + 2\bar{n}) \quad (47)$$

for the squeezed-state channel. As we see noise beat usefulness of squeezed states, and for optimal squeezing parameter we have (44).

Now about general case for the homodyne measurement. We get

$$I(\alpha : \gamma_1/\beta) = \frac{1}{2} \ln \left(1 + \frac{(\bar{n}_1 - \text{sh}^2 r_1)u_1^2}{\frac{1}{4}(u_1^2 e^{-2r_1} + u_2^2 e^{2r_1}) + \frac{1}{4}(v_1^2 e^{-2r_2} + v_2^2 e^{2r_2}) + \frac{1}{4}\psi} \right) \quad (48)$$

$$I(\alpha \otimes \beta : \gamma_1) = \frac{1}{2} \ln \left(1 + \frac{(\bar{n}_1 - \text{sh}^2 r_1) u_1^2 + (\bar{n}_2 - \text{sh}^2 r_2) v_1^2 + \frac{1}{4}(u_1^2 e^{-2r_1} + u_2^2 e^{2r_1}) + \frac{1}{4}(v_1^2 e^{-2r_2} + v_2^2 e^{2r_2}) + \frac{1}{4}\Psi}{\frac{1}{4}\Psi} \right) \quad (49)$$

Where

$$\begin{aligned} u_1 &= e^{-\gamma t/2} (\epsilon \cos(\lambda_1 t) + (1 - \epsilon) \cos(\lambda_2 t)) \\ u_2 &= -e^{-\gamma t/2} (\epsilon \sin(\lambda_1 t) + (1 - \epsilon) \sin(\lambda_2 t)) \\ v_1 &= -e^{-\gamma t/2} \sqrt{\epsilon(1 - \epsilon)} (\cos(\lambda_1 t) - \cos(\lambda_2 t)) \\ v_2 &= e^{-\gamma t/2} \sqrt{\epsilon(1 - \epsilon)} (\sin(\lambda_1 t) - \sin(\lambda_2 t)) \end{aligned} \quad (50)$$

We can maximize by r_1, r_2 the information measure $I(\alpha : \gamma/\beta)$ (as we know it is information transmitted by user 1), and after this information measure $I(\alpha \otimes \beta : \gamma) - I(\alpha : \gamma/\beta)$ for user 2. The analysis show that situation with one-terminal channel is conserved: there are optimal r_1, r_2 which tend to zero for large t . The second mode introduce additional source for noise.

4 Conclusion.

We have considered the noisy dual-access quantum-mechanical channel, and compute capacities of this channel. The general theorems are proved which connect capacities of the channel with some functions from statistical mechanics. It is shown that each user acts as noise and can significantly reduce the capacity of the other user. After this the information transfer by coherent and squeezed states of an electromagnetic field is discussed. The squeezed states loss optimality under action of the noise and the optimal squeezing parameter tends to zero when time tends to infinity.

References

1. C.M. Caves, P.D. Drummond, Rev.Mod.Phys., **66**, 481, (1994).
2. Gordon J.P., Proc. IRE, **50**, 1898, (1962). Lebedev D.S., and L.B. Levitin, Dokl. Akad. Nauk SSSR, **149**, 1299, (1963), (Sov. Phys. Dokl. **8**, 377, (1963))
3. A.S. Holevo, e-print, quant-ph, 96011023.
4. A.S. Holevo, Problems of Information Transmission, **9**, 3, (1973).
5. C.W. Helstrom, *Quantum detection and estimation theory*. Academic Press, 1976.
6. Yuen H.P., Ozawa M., Phys. Rev. Lett. **70**, 363, (1993).
7. B.Schumacher, Phys. Rev. A **54**, 2614, (1996).
8. P. Hausladen et al., Phys.Rev. A, **54**, 1869, (1996).
9. R. L. Stratonovich, *Information Theory*, Moscow, Nauka 1975.
10. A. El Gammal, T. Cover, Proc. IEEE, **68**, 1466, (1980).
11. K.Kraus, Ann. Phys., **64**, 311, (1971).
12. G.Lindblad, Commun. Math. Phys., **40**, 147, (1975).
13. M.Ohya, Rep. Math. Phys., **27**, 19, (1989).
14. H.Umegaki, Kodai Math. Sem. Rep., **14**, 59, (1962).
15. K. Luissel, *Radiation and nise in quantum electronics*. McGrow-Hill Book Company, 1964.
16. M. Lax, Phys.Rev., **145**, 110, (1966).

Capacities of Quantum Channels and Quantum Coherent Information

Michael Westmoreland¹ and Benjamin Schumacher²

¹ Department of Mathematics, Denison University, Granville, OH 43023

² Department of Physics and Astronomy, Kenyon College, Gambier, OH 43022

Abstract. We derive relation between a quantum channel's capacity to convey classical information and its ability to convey quantum information. We also show that these properties of a quantum channel are related to the channel's ability to convey quantum coherent information.

1 Introduction

A quantum communication channel can be used to perform a variety of tasks, including:

1. Conveying classical information from a sender to a receiver.
2. Conveying quantum information (including quantum entanglement) from a sender to a receiver.

Each of these tasks can be performed in the presence of noise. Indeed, in quantum cryptography the noise is of central importance in revealing the activity of an eavesdropper.

A central concern in the analysis of any noisy communications channel is the channel's *capacity*: the maximum rate at which information (classical or quantum) can be reliably transmitted through the channel. When we use quantum systems to convey information we confront concerns which are not present in classical channels. A chief concern is the measurement performed by the receiver in order to extract the information.

Measurement of the received message is not particular to quantum channels; a receiver using a classical channel must also measure the received message. Counting the numbers of dots and dashes in a Morse codeword is a measurement. What *is* a particular concern for the user of a quantum channel is that a measurement result may be ambiguous *even if the quantum channel is noiseless*. For example, assume that Alice, the sender and Bob, the receiver, agree to use vertically polarized photons to represents a “1” and horizontally polarized photons to represent a “0”. We also assume that Bob measures for photons polarized along an axis that is 45 degrees from horizontal. Quantum mechanics tells us that, in this case, Bob could not tell if Alice sent a “1” or a “0”. This is true even if Alice's photon reaches Bob without ant distortion in its angle of polarization. One might object (quite rightfully in this case) that Alice and Bob settled on a particularly silly measurement to use in reading the signals.

If the letter states (photon polarizations here) are not orthogonal then there is no measurement Bob can perform that will perfectly distinguish them. This is a first difference between classical and quantum channels: measurements are, in general, ambiguous, even if no noise is present.

One might suggest that if Bob cannot use a single measurement to decode Alice's signal then Bob should perform several measurements on each photon. Quantum mechanics prevents implementation of such a scheme: a measurement of a quantum system fundamentally changes the properties of that system. We return to our example where Alice and Bob used vertical and horizontal polarizations and a measurement at 45 degrees. After Bob has measured a given photon, it is in a pure state of 45 degree polarization. That is, if a second polarizer is set up immediately after the first with the same (45 degree) polarization then the photon will pass it with certainty. If the polarizer is set at any other angle, then there is a nonzero chance that the particle will not pass. Indeed, even if Alice dispatches the photon with vertical polarization and if it passes Bob's first polarizer, it will then have only a 50% chance of passing a subsequent vertical polarizer. This is a second difference between classical and quantum channels: measurements of the received signals fundamentally change the signal.

One might further suggest that if Bob's measurements change the state of the transmitted systems, then Bob should make several copies of each signal state. He could then perform a single measurement on each copy and thereby extract Alice's message. Once again, quantum mechanics prevents such a solution. The problem here is that general quantum states can not be perfectly copied. This is the content of the no-cloning theorem of Wootters and Zurek [3]. This is a third difference between quantum and classical channels: quantum states cannot be cloned.

It should be noted that the no-cloning theorem does not imply that Alice can not make multiple copies of the states she is sending. All that she needs to do is prepare multiple systems in the same way; this is not cloning. If Alice and Bob decide to do this to increase the reliability of their channel it would decrease the capacity of the channel. This is because capacity is the *rate* at which information is transmitted per letter state. If Alice send more letter states, the capacity of the channel is reduced.

Given what has been said to this point, one might conclude that even finding the classical capacity of a quantum channel is problematic at best. In fact, we do know the classical capacity, even for noisy quantum channels. This result is presented in Section II.

To this point, we have been concerned with the classical capacity of a quantum channel. One can also analyze the *quantum* capacity of a quantum communications channel. This is the ability of the channel to faithfully transmit a quantum state from one system to another. The capacity of a noiseless quantum channel is known [11] but the case of the noisy channel is still under analysis even though some progress has been made [20]. Section III presents the concept of coherent quantum information [18] and some of its properties. Section III concludes with an analysis showing the relation between the classical capacity of

a channel, the coherent information conveyed by that channel and the quantum capacity of the channel.

2 Classical Capacity of a Noisy Quantum Channel

Suppose Alice wishes to convey classical information to Bob by using a quantum system Q as a communication channel. Alice prepares the channel in one of various quantum states W_x with *a priori* probabilities p_x . Bob makes a measurement on the system Q , and from its result he tries to infer which state Alice prepared. A theorem stated by Gordon [1] and Levitin [5], first proved by Kholevo [6], gives an upper bound to the amount of information that Bob can obtain about Alice's signal. If $W = \sum_x p_x W_x$ is the density operator describing the ensemble of Alice's signals, then the mutual information $H(X : Y)$ between Alice's input X and Bob's output Y is bounded by

$$H(X : Y) \leq H(W) - \sum_x p_x H(W_x), \quad (1)$$

where $H(W) = -\text{Tr } W \log W$, the von Neumann entropy of the density operator W . The upper bound in Equation (1) is in general a weak one, in that Bob may not be able to choose an observable that gives him an amount of information near to the upper bound [7].

Suppose that Alice employs signal states W_x that are *mixed* states. Then can Alice and Bob find a choice of code and decoding observable so that the general Levitin - Kholevo bound (Equation (1)) can be approached arbitrarily closely? In this paper, we show that the answer to this question is "yes". That is, we prove the following result:

Theorem. Suppose we have letter states W_x with *a priori* probabilities p_x , and let

$$\chi = H(W) - \sum_x p_x H(W_x).$$

Fix $\epsilon, \delta > 0$. Then for sufficiently large L , there exist a code (whose codewords are strings of L letters) and a decoding observable such that the information carried per letter is at least $\chi - \delta$ and the probability of error $P_E < \epsilon$.

The proof employs an average over randomly generated codes to establish the existence of a satisfactory code. (If the average probability of error is small for an ensemble of codes, the ensemble must contain specific codes with small probability of error.) We also use a similar prescription for Bob's decoding observable. The chief refinement in the proof presented here is the enforcement of stronger "typicality" conditions on various quantities associated with the channel.

The mixed states W_x may be thought of as the outputs of a *noisy* quantum channel. Thus, our main result will enable us to draw conclusions about the classical information capacity of a noisy quantum channel.

We have shown that it is possible to send information at any rate up to χ bits per letter with arbitrarily low probability of error. The capacity of a channel is defined as the maximum information per letter that may be sent through the channel with P_E arbitrarily small. Thus, χ provides a lower bound to the capacity of the quantum channel.

Classical information theory together with the Levitin - Kholevo Theorem also allows us to use χ to establish an *upper bound* for the capacity of the channel. Suppose X represents Alice's input and Y represents Bob's decoding measurement outcome. Then the Fano inequality [10] states that

$$-P_E \log P_E - (1 - P_E) \log(1 - P_E) + P_E \log(N_X - 1) \geq H(X|Y) \quad (2)$$

where P_E is the probability of error and N_X is the number of possible values of X . $H(X|Y)$ is the conditional Shannon entropy of X given Y —that is, the entropy of the conditional distribution $p(x|y)$, averaged over the various values of y . It is related to the mutual information $H(X : Y)$ by

$$H(X|Y) = H(X) - H(X : Y). \quad (3)$$

In the channel, Alice uses some signal states ρ_a with probabilities P_a . Levitin - Kholevo Theorem places an upper bound on the mutual information $H(X : Y)$:

$$H(X : Y) \leq H(\rho) - \sum_a P_a H(\rho_a).$$

(Note that, if the channel used by Alice and Bob consists of L letters used independently, then the Levitin - Kholevo bound is just $L\chi$, where χ is the Levitin - Kholevo bound for a single letter.) If the Alice's input X has an entropy $H(X)$ that exceeds $H(\rho) - \sum_a P_a H(\rho_a)$, then $H(X|Y) > 0$ and it will not be possible to make the probability of error P_E arbitrarily small.

Suppose we fix an alphabet $\Gamma = \{W_x\}$ of letter states W_x , and require that Alice use codewords a that are length- L strings of these letter states: $a = x_1 \dots x_L$. Then the probability distribution P_a yields marginal probability distributions $p(x_1), \dots, p(x_L)$ and average density operators W_1, \dots, W_L for the L different letters. It follows that

$$\begin{aligned} H(\rho) - \sum_a P_a H(\rho_a) &\leq \left(H(W_1) - \sum_{x_1} p(x_1) H(W_{x_1}) \right) \\ &\quad + \dots + \left(H(W_L) - \sum_{x_L} p(x_L) H(W_{x_L}) \right) \end{aligned} \quad (4)$$

where we have used the subadditivity of the entropy $H(\rho)$. We might write this as

$$\chi^{(L)} \leq \chi_1 + \dots + \chi_L \quad (5)$$

where $\chi^{(L)}$ represents the Levitin - Kholevo bound for the ensemble of codewords of length L , and χ_1, \dots, χ_L represent Levitin - Kholevo bounds for the individual letter ensembles.

We define the *fixed-alphabet capacity* C_Γ to be

$$C_\Gamma = \sup_{p(x)} \chi \quad (6)$$

where $p(x)$ is the probability distribution over the letters states in Γ and χ is the single-letter Levitin - Kholevo bound. This quantity represents the maximum information rate per letter that Alice can send to Bob with arbitrarily low probability of error.

This claim follows directly from our results so far. Suppose Alice uses codewords of length L . Then $\chi^{(L)} \leq L C_\Gamma$; and by the above argument, if Alice attempts to send more than $L C_\Gamma$ bits using these codewords then the probability of error will not be arbitrarily small. Conversely, we can choose the letter probabilities so that χ is as close as required to C_Γ , and we have previously shown that a suitable choice of code and decoding observable can convey up to χ bits per letter with arbitrarily low P_E . Thus, the capacity C_Γ cannot be exceeded but can be approached arbitrarily closely.

The mixed states W_x used in our alphabet are the states available to Bob for decoding. They may in fact not be the original states of the channel Q chosen by Alice. In the interval between Alice's encoding and Bob's decoding, the system Q may have undergone unitary internal evolution (which Bob can correct by a suitable choice of "rotated" decoding observable) and interaction with the external environment (which Bob cannot in general correct).

The most general description of the evolution of a quantum system Q interacting with an environment is provided by a trace-preserving completely positive linear map on the set of density operators of Q [14]. Such a map is described by a superoperator \mathcal{E} :

$$\rho \longrightarrow \rho' = \mathcal{E}(\rho), \quad (7)$$

where ρ is the initial state of the system and ρ' is the final state. The superoperator \mathcal{E} acts linearly, so that a convex combination of input states yields a convex combination of output states. This description clearly includes unitary evolution of Q as a special case, but it also can account for interaction with the environment.

A noisy quantum channel is defined by a superoperator \mathcal{E} that describes the evolution of each letter as it is transmitted from Alice to Bob. We assume that the channel is memoryless—i.e., that the evolution of each letter is independent. This means, among other things, that a product state of several input letters will evolve into a product state output.

Alice's basic problem is to use input states w_x so that the output states $W_x = \mathcal{E}(w_x)$ can be distinguished by Bob. If Alice has a fixed alphabet $\{w_x\}$ of input states, then the maximum achievable information rate per letter is still given by our fixed-alphabet capacity C_Γ , where Γ is the alphabet of *output* states.

Now suppose that Alice is allowed to choose her input states in order to maximize the information conveyed to Bob over the noisy quantum channel, subject to the constraint that Alice must transmit codewords which are represented by

product states of the letters. This *almost* reduces to the fixed-alphabet problem, where the fixed alphabet \mathcal{I} now includes all of the possible output states of the channel. The maximum over probability distributions is now a maximum over all input ensembles of states chosen by Alice.

We say that this problem *almost* reduces to the fixed alphabet problem in that the argument that χ is an upper bound of the capacity must be modified in this case. Recall from the previous section that we applied the classical Fano inequality to show that if Alice attempts to send information at a rate exceeding χ then the probability of error cannot be made arbitrarily small. If we attempt to use the same argument in the present case then the Fano inequality does not help us for at least two reasons. First, the number of possible input states N_x is unbounded. Second, we do not have a characterization of $H(X|Y)$ that allows us to compare it with N_x . Thus we will modify the Fano inequality to understand the behavior of the probability of error in the present case.

We first note that the probability of “getting it right”

$$1 - P_E = \frac{1}{N} \sum_{\alpha k} p_{k|\alpha} |\langle \tilde{\mu}_{\alpha k} | s_{\alpha k} \rangle|^2 \quad (8)$$

is linear in the elements of the POM. Thus the probability of error, P_E is a convex function on the elements of the POM. We may modify the proof of a result of Davies (Theorem 3 of [17]) to show that the convex function P_E is minimized by a POM having no more than d^2 elements, where d is the dimension of the support of the POM. Thus, the probability of error is minimized by a decision scheme in which at most d^2 of the inputs are identified by the decision scheme. Let us denote the output of such a scheme by Y_{min} . Fano’s inequality gives us that

$$-P_E \log P_E - (1 - P_E) \log(1 - P_E) + P_E \log(d^2 - 1) \geq H(X|Y_{min}). \quad (9)$$

Note that

$$H(X|Y_m) = H(X) - H(X : Y_{min}) \quad (10)$$

$$\geq H(X) - \chi, \quad (11)$$

so that we conclude

$$-P_E \log P_E - (1 - P_E) \log(1 - P_E) + P_E \log(d^2 - 1) \geq H(X) - \chi. \quad (12)$$

Note that this is a relation between the minimum probability of error and a quantity $(H(X) - \chi)$ which does not depend on the particular decision scheme. We see that if Alice attempts to send information at a rate $H(X)$ in excess of χ then the probability of error can not be made arbitrarily small.

We now turn to a demonstration that this rate can be achieved. Alice wishes to choose a set of input states w_x (together with input probabilities p_x) so that χ is maximized for the output states W_x . We next show that Alice can do no

better than choose the input states w_x to be pure. Let a set of (possibly mixed) input states w_x be given along with their a priori probabilities, and let

$$W = \sum_x p_x W_x = \sum_x p_x \mathcal{E}(w_x) \quad (13)$$

be the average output state. Then

$$\chi = H(W) - \sum_x p_x H(\mathcal{E}(w_x)). \quad (14)$$

Construct a new set of pure state inputs by resolving each mixed state input into a convex combination of pure states:

$$w_x = \lambda_{xk} |\psi_{xk}\rangle \langle \psi_{xk}|. \quad (15)$$

We will use the state $|\psi_{xk}\rangle$ with probability $p_{xk} = p_x \lambda_{xk}$. By linearity,

$$W_x = \mathcal{E}(w_x) = \sum_k \lambda_{xk} \mathcal{E}(|\psi_{xk}\rangle \langle \psi_{xk}|), \quad (16)$$

so that the average output state is still W , as before. By the convexity of the von Neumann entropy,

$$H(W_x) \geq \sum_k \lambda_{xk} H(\mathcal{E}(|\psi_{xk}\rangle \langle \psi_{xk}|)). \quad (17)$$

It follows that

$$\begin{aligned} \chi' &= H(W) - \sum_{xk} p_{xk} H(\mathcal{E}(|\psi_{xk}\rangle \langle \psi_{xk}|)) \\ &\geq H(W) - \sum_x p_x H(\mathcal{E}(w_x)) = \chi. \end{aligned} \quad (18)$$

In other words, for any ensemble of mixed input states, we can find an ensemble of pure input states whose output states have a χ at least as great. The optimal inputs for the noisy quantum channel are pure states.

To sum up, if Alice is required to use product states to represent her code-words, then the capacity $C^{(1)}$ of the noisy quantum channel is

$$C^{(1)} = \max \chi \quad (19)$$

where χ is the Levitin - Kholevo bound for the output states of the channel, and the maximum is taken over all ensembles of pure state inputs. Alice can reliably transmit information to Bob at any rate below $C^{(1)}$. We will refer to $C^{(1)}$ as the product state capacity. The superscript (1) reminds us that Alice is required to use the multiple available copies of the channel *one at a time*, coding her messages into product states.

3 Coherent Quantum Information and Quantum Capacity

The entropy exchange S_e measures the amount of information that is exchanged between the system Q and the environment E during their interaction. If the environment is initially in a pure state, the entropy exchange is just the environment's entropy after the interaction—i.e., $S_e = S(\rho^{E'})$, where $\rho^{E'}$ is the final state of E . (The entropy here is just the ordinary von Neumann entropy of a density operator, $S(\rho) = -\text{Tr } \rho \log \rho$.) The entropy exchange is entirely determined by the initial state ρ^Q of Q and the channel dynamics superoperator \mathcal{E}^Q ; that is, the entropy exchange is a property “intrinsic” to Q and its dynamics.

The coherent information I_e , introduced in [18], is given by

$$I_e = S(\rho^{Q'}) - S_e. \quad (20)$$

The coherent information has many properties that suggest it as the proper measure of the quantum information conveyed from Alice to Bob by the channel. For example, I_e can never be increased by quantum data processing performed by Bob on the channel output, and perfect quantum error correction of the channel output is possible for Bob if and only if no coherent information is lost in the channel [18]. Finally, the coherent information seems to be related to the capacity of a quantum channel to convey quantum states with high fidelity [20].

Alice might be using the channel to send classical information to Bob. Alice prepares Q in one of a set of possible “signal states” ρ_k^Q , which are used by Alice with *a priori* probabilities p_k . The average state ρ^Q is given by

$$\rho^Q = \sum_k p_k \rho_k^Q. \quad (21)$$

Bob receives the k th signal as $\rho_k^{Q'} = \mathcal{E}^Q(\rho_k^Q)$. Because the superoperator is linear, the average received state is

$$\rho^{Q'} = \sum_k p_k \mathcal{E}^Q(\rho_k^Q) = \mathcal{E}^Q(\rho^Q). \quad (22)$$

Bob attempts to decode Alice's message (that is, to identify which signal state was chosen by Alice) by measuring some *decoding observable* on his received system Q' .

The amount of classical information conveyed from Alice to Bob, which we will denote H_{Bob} , is governed by the quantity $\chi^{Q'}$, defined by

$$\chi^{Q'} = S(\rho^{Q'}) - \sum_k p_k S(\rho_k^{Q'}). \quad (23)$$

This quantity is significant in two ways:

- $H_{Bob} \leq \chi^{Q'}$, regardless of the decoding observable chosen [21,22].
- H_{Bob} can be made as close as desired to $\chi^{Q'}$ by a suitable choice of code and decoding observable. To make H_{Bob} near $\chi^{Q'}$, Alice must in general use the channel many times and employ code words composed of many signals; Bob must perform his decoding measurement on entire code words. The net result is that the channel is used N times to send up to $N\chi^{Q'}$ bits of classical information reliably [2].

In short, $\chi^{Q'}$ represents an upper bound on the classical information conveyed from Alice to Bob, an upper bound that may be approached arbitrarily closely if Alice and Bob use the channel efficiently.

If this general picture is used to describe a noisy quantum channel, then we need to account for the information that is passed to the environment. Recall that the evolution superoperator \mathcal{E}^Q describes all of the effects of the channel; or, to put it another way, all of properties of the link between Alice and Bob are contained in the interaction operator U^{QE} . The information passed to the environment H_E will be limited by

$$\chi^{E'} = S(\rho^{E'}) - \sum_k p_k S(\rho_k^{E'}). \quad (24)$$

Assume that the states of Q initially prepared by Alice are pure states $|\phi_k^Q\rangle$; also recall that the environment E can be presumed to begin in a pure state $|0^E\rangle$. After Q and E interact unitarily, the joint state $|\Psi_k^{QE'}\rangle = U^{QE} |\phi_k^Q\rangle \otimes |0^E\rangle$ will also be a pure state, generally an entangled one. The subsystem states, described by density operators

$$\begin{aligned} \rho_k^{Q'} &= \text{Tr}_E |\Psi_k^{QE'}\rangle \langle \Psi_k^{QE'}| \\ \rho_k^{E'} &= \text{Tr}_Q |\Psi_k^{QE'}\rangle \langle \Psi_k^{QE'}|, \end{aligned} \quad (25)$$

will have exactly the same non-zero eigenvalues, so that $S(\rho_k^{Q'}) = S(\rho_k^{E'})$. Therefore

$$\begin{aligned} I^Q &= S(\rho^{Q'}) - S_e \\ &= S(\rho^{Q'}) - S(\rho^{E'}) \\ &= S(\rho^{Q'}) - \sum_k p_k S(\rho_k^{Q'}) - S(\rho^{E'}) + \sum_k p_k S(\rho_k^{E'}) \\ I^Q &= \chi^{Q'} - \chi^{E'}. \end{aligned} \quad (26)$$

It is interesting to note that, although both $\chi^{Q'}$ and $\chi^{E'}$ depend on the choice of pure state inputs for the channel Q , the difference $\chi^{Q'} - \chi^{E'}$ depends only on the overall density operator ρ^Q for the inputs.

It is shown in [18] that perfect error correction is possible if and only if the coherent information of the channel equals the entropy of the input state. The quantity D_e is defined [24] as follows:

$$D_E = S(\rho^Q) - I^Q. \quad (27)$$

We can thus say that perfect error correction is possible if and only if $D_e = 0$.

Subtracting each side of equation (26) from the entropy of the input state yields:

$$S(\rho^Q) - I^Q = D_e = S(\rho^Q) - \chi^{Q'} + \chi^{E'}. \quad (28)$$

Recall that χ^Q is maximized when Alice uses pure state to encode her messages. In that case we have $\chi^Q = S(\rho^Q)$, so equation (28) takes the form

$$D_e = \chi^Q - \chi^{Q'} + \chi^{E'}. \quad (29)$$

This equation is quite informative. It implies that conveying quantum information perfectly depends on two tasks: Maximization of classical capacity and zero entropy loss to the environment. This implies a strong connection between the quantum capacity of a quantum channel and its classical capacity.

We would like to thank W. K. Wootters and M. A. Nielsen for helpful conversations and suggestions.

References

1. J. P. Gordon, in *Quantum Electronics and Coherent Light, Proceedings of the International School of Physics "Enrico Fermi," Course XXXI*, edited by P. A. Miles (Academic, New York, 1964), pp. 156-181.
2. P. Hausladen, R. Jozsa, B. Schumacher, M. D. Westmoreland and W. K. Wootters, *Phys. Rev. A* **54**, 1869 (1996). B. Schumacher and M. D. Westmoreland, *Phys. Rev. A* **56**, 131 (1997). A. S. Holevo, *IEEE Trans. Inf. Theory* (to be published).
3. W.K. Wootters and W.H. Zurek, *Nature* **299**, 802 (1982)
4. B. Schumacher and M. Westmoreland, "Coherent quantum information and quantum privacy", submitted to *Phys. Rev. Lett.*
5. L. B. Levitin, "On the quantum measure of the amount of information," in *Proceedings of the IV National Conference on Information Theory*, Tashkent, 1969, pp. 111-115 (in Russian); "Information Theory for Quantum Systems," in *Information, Complexity, and Control in Quantum Physics*, edited by A. Blaqui re, S. Diner, and G. Lochak (Springer, Vienna, 1987).
6. A. S. Kholevo, *Probl. Inform. Transmission* **9**, 177 (1973) (translated from *Problemy Peredachi Informatsii*).
7. C. A. Fuchs and C. M. Caves, *Phys. Rev. Lett.* **73**, 3047 (1994).
8. P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, *Phys. Rev. A* **54**, 1869 (1996).
9. A. S. Kholevo, to appear in *IEEE Transactions on Information Theory*.
10. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
11. B. Schumacher, *Phys. Rev. A* **51**, 2738 (1995).

12. B. Schumacher and R. Jozsa, *J. Mod. Opt.* **41**, 2343 (1994).
13. A. S. Kholevo *Probl. Inform. Transmission* **15**, 3 (1979) (translated from *Problemy Peredachi Informatsii*).
14. K. Hellwig and K. Kraus, *Communications in Mathematical Physics* **16**, 142 (1970). M.-D. Choi, *Linear Algebra and its Applications* **10**, 285 (1975). K. Kraus, *States Effects and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, Berlin, 1983).
15. C. H. Bennett, C. A. Fuchs, and J. Smolin, to appear in *Proceedings of the 3rd International Workshop on Quantum Communication and Measurement*.
16. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379, 623 (1948).
17. E. B. Davies, *IEEE Transactions on Information Theory* **IT-24**, 596 (1978).
18. B. Schumacher and M. A. Nielsen, *Phys. Rev. A* **54**, 2629 (1996).
19. B. Schumacher, *Phys. Rev. A* **54**, 2614 (1996).
20. S. Lloyd, *Phys. Rev. A* **55**, 1613 (1997). H. Barnum, M. A. Nielsen and B. Schumacher, Report No. quant-ph/9702049.
21. A. S. Kholevo, *Probl. Peredachi Inf.* **9**, 3 (1973) [*Probl. Inf. Transm. (USSR)* **9**, 110 (1973)].
22. B. Schumacher, M. D. Westmoreland and W. K. Wootters, *Phys. Rev. Lett.* **76**, 3452 (1996).
23. I. Csiszár and J. Körner, *IEEE Transactions on Information Theory* **24**, 339 (1978). U. M. Maurer, *IEEE Transactions on Information Theory* **39**, 733 (1993).
24. B. Schumacher, "Entropy exchange and coherent quantum information" in *Proceedings of the Fourth Workshop on Physics and Computation*; edited by T. Toffoli, M. Biafore, and J. Leao (New England Complex Systems Institute, Boston, 1996).

Strengthened Lindblad Inequality: Applications in Non-equilibrium Thermodynamics and Quantum Information Theory

D.B. Saakian and A.E. Allahverdyan

Yerevan Physics Institute
Alikhanian Brothers St.2, Yerevan 375036, Armenia

Abstract. Strengthened Lindblad inequality has been proved. We have applied this results for proving a generalized H -theorem in non equilibrium thermodynamics. Information processing also can be considered as some thermodynamically process. From this point of view we have proved strengthened data processing inequality in quantum information theory.

There are close connection between statistical thermodynamics and information theory [2,4]. It is well known that physical ideas played an important role as sources of information theory [2]. In the other hand, the concept of information is crucial for understanding some important physical problems such as Maxwell "demon" [4] or general problem of quantum correlation between two quantum systems [5].

In this paper we concentrated on the two connected problems. There are the H -theorem problem in non equilibrium thermodynamics and problem of quantum data processing in quantum information theory.

Suppose that a quantum system is described by density matrix $\rho(t)$ at the moment t . In the general case evolution of the non equilibrium system in markoffian regime is described by some general quantum evolution operator

$$\hat{S}(t', t)\rho(t) = \rho(t') \quad (1)$$

This most general deterministic quantum evolution operator \hat{S} has kraussian representation

$$\hat{S}\rho = \sum_{\mu} A_{\mu}^{\dagger} \rho A_{\mu}, \quad \sum_{\mu} A_{\mu} A_{\mu}^{\dagger} = \hat{1}. \quad (2)$$

This operators must be linear, completely positive and trace-preserving [6,10]. The (2) contains unitary evolution, nonselective measurement, partial trace, et all. If the evolution of the system is in the stationary markoffian regime then

$$\hat{S}(t', t) = \hat{S}(t' - t) \quad (3)$$

Stationary markoffian regime is reasonable conjecture if the system is not far from equilibrium [11,3].

When the system is open we need some values, which must characterize the degree of nonunitarity for evolution of our system. The usual characteristics for this case is entropy. This quantity was introduced in quantum statistical physics by J. von Newmann.

$$S(\rho) = -\text{tr} \rho \ln \rho \quad (4)$$

As we will see later for some more general problems we need a more general quantity. Quantum relative entropy between two density matrices ρ_1, ρ_2 is defined as follows

$$S(\rho_1 || \rho_2) = \text{tr}(\rho_1 \log \rho_1 - \rho_1 \log \rho_2). \quad (5)$$

This positive quantity was introduced by Umegaki [12] and characterizes the degree of 'closeness' of density matrices ρ_1, ρ_2 . The properties of quantum relative information were reviewed by M.Ohya [11]. Here only two basic properties are mentioned.

$$S(\rho_1 || \rho_2) \geq S(\hat{S}\rho_1 || \hat{S}\rho_2). \quad (6)$$

$$S(\lambda\rho_1 + (1-\lambda)\sigma_1 || \lambda\rho_2 + (1-\lambda)\sigma_2) \leq \lambda S(\rho_1 || \sigma_1) + (1-\lambda)S(\rho_2 || \sigma_2). \quad (7)$$

Where $0 \leq \lambda \leq 1$. The first inequality was proved by Lindblad [13].

When we have a usual H -theorem? If (3) has a stationary density matrix ρ_{st} and this ρ_{st} equal unit matrix (up to unessential normalization constant), then

$$\hat{S}(t' - t)1 = 1 \quad (8)$$

for any $t' - t$. The condition $\rho_{st} \sim 1$ is a usual microcanonical distribution in equilibrium statistical physics. Then from (6) we have

$$S(\hat{S}\rho) \geq S(\rho) \quad (9)$$

In general case (when $\rho_{st} \not\sim 1$, it is usual case in the theory of open systems) entropy of von Newmann is not monotonically increasing function with time and for some open system can exhibit periodic behavior [14]. In general case the author of [1] proposed to use as measure of nonunitarity of evolution the relative entropy between ρ_{st}, ρ . Indeed, if we define the entropy of our system at the time t as

$$-S(\rho(t) || \rho_{st}) \quad (10)$$

then from (6) we see that for any type of evolution (10) is monotonic function with time. Furthermore (10) has all physically important properties of usual entropy (8): it is positive, additive and convex [11].

Now the following question arises. Can we generalize the (6) without any restrictions? If the answer is yes, then we can prove with this result more general H -theorem. Let us assume in formula (6) that

$$\hat{S} = c\hat{C}_1 + (1-c)\hat{C}_2, \quad (11)$$

where \hat{C}_1 is defined by kraussian representation $A_\mu = |\mu\rangle\langle 0|, \langle \mu|\hat{\mu}\rangle = \delta_{\mu\hat{\mu}}, \langle 0|0\rangle = 1, 0 \leq c \leq 1$. In other words for any operator $\rho: \hat{C}_1\rho = |0\rangle\langle 0|$. Now from (6), (7) we get

$$\begin{aligned} S(\hat{S}\rho||\hat{S}\sigma) &= S(c\hat{C}_1\rho + (1-c)\hat{C}_2\rho||c\hat{C}_1\sigma + (1-c)\hat{C}_2\sigma) \\ &\leq cS(\hat{C}_1\rho||\hat{C}_1\sigma) + (1-c)S(\hat{C}_2\rho||\hat{C}_2\sigma) \leq (1-c)S(\rho||\sigma). \end{aligned} \quad (12)$$

We see that if \hat{S} is represented in the form (11) the ordinary Lindblad inequality can be strengthened.

Now we need some general results from theory of linear operators [17]. Let two hermitian operators A and B have the spectrums $a_1 \leq \dots \leq a_n$, $b_1 \leq \dots \leq b_n$. For the spectrum $c_1 \leq \dots \leq c_n$ of the operator $C = A + B$ we have

$$a_1 + b_k \leq c_k \leq b_k + a_n, \quad b_1 + a_k \leq c_k \leq a_k + b_n. \quad (13)$$

where $k = 1, \dots, n$. If

$$\begin{aligned} \rho' &= \hat{S}\rho = c\hat{C}_1\rho + (1-c)\hat{C}_2\rho \\ &= c|0\rangle\langle 0| + (1-c)\sigma, \end{aligned} \quad (14)$$

and $\rho'_1 \leq \dots \leq \rho'_n$, $\sigma_1 \leq \dots \leq \sigma_n$ are the spectrums of ρ' , σ then we have

$$\begin{aligned} \rho'_1 - c &\leq \sigma_1(1-c) \leq \min(\rho'_1, \rho'_n - c), \\ \max(\rho'_1, \rho'_k - c) &\leq \sigma_k(1-c) \leq \rho'_k, \end{aligned} \quad (15)$$

where $k = 2, \dots, n$. We define $c(\hat{S}, \rho)$ as the minimal eigenvalue of ρ' and $c(\hat{S}) = \min_\rho c(\hat{S}, \rho)$ where minimization is taken by all density matrices for the fixed Hilbert space. With the well known results of operator theory [17] we can write

$$c(\hat{S}) = \min_\rho \min_{\langle \psi | \psi \rangle = 1} \langle \psi | \hat{S} \rho | \psi \rangle, \quad (16)$$

where the second minimization is taken by all normal vectors in the Hilbert space. For any density matrix ρ we get to the formula (11) where c is defined in (16) and \hat{C}_2 is some general evolution operator. Now from (11), (12), (16) we get the strengthened Lindblad inequality

$$(1-c)S(\rho_1||\rho_2) \geq S(\hat{S}\rho_1||\hat{S}\rho_2). \quad (17)$$

The equations (16), (17) are our general results. Of course there are many evolution operators \hat{S} with $c(\hat{S}) = 0$ but later we shall show that our results can be nontrivial because for some simple but physically important case $c(\hat{S})$ is nonzero. From (16), (17) we immediately get to strengthened H -theorems.

Now about application of this result in quantum data processing.

Quantum information theory is a new field with potential applications for the conceptual foundation of quantum mechanics. It appears to be the basis for a proper understanding of the emerging fields of quantum computation, communication and cryptography (see [6] for references). Quantum information theory concerned with quantum bits (qubits) rather than bits. Qubits can exist in superposition or entanglement states with other qubits, a notion completely inaccessible for classical mechanics. More general, quantum information theory contains two distinct types of problem. The first type describes transmission of

classical information through a quantum channel (the channel can be noisy or noiseless). In such scheme bits encoded as some quantum states and only this states or its tensor products are transmitted. In the second case arbitrary superposition of this states or entanglement states are transmitted. In the first case the problems can be solved by methods of classical information theory, but in the second case new physical representations are needed.

Mutual information is the most important ingredient of information theory. In classical theory this value was introduced by C.Shannon [16]. The mutual information between two ensembles of random variables X, Y (for example this ensembles can be input and output for a noisy channel)

$$I(X, Y) = H(Y) - H(Y/X), \quad (18)$$

is the decrease of the entropy of X due to the knowledge about Y , and conversely with interchanging X and Y . Here $H(Y)$ and $H(Y/X)$ are Shannon entropy and mutual entropy [16].

Mutual information in the quantum case must take into account the specific character of the quantum information as it is described above. The reasonable definition of this quantity was first introduced by S.Lloyd [9], and independently by B.Schumacher and M.P.Nielsen [7]. Suppose a quantum system with density matrix

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad \sum_i p_i = 1. \quad (19)$$

We only assume that $\langle\psi_i|\psi_i\rangle = 1$ and the states may be nonorthogonal. The noisy quantum channel can be described by some general quantum evaluation operator \hat{S} .

As follows from definition of quantum information transmission, a possible distortion of entanglement of ρ must be taken into account. In other words definition of mutual quantum information must contain the possible distortion of relative phases of quantum ensemble $\{|\psi_i\rangle\}$. Mutual quantum information is defined as [9, 7]

$$I(\rho; \hat{S}) = S(\hat{S}\rho) - S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|)), \quad (20)$$

$$\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|) = \sum_{i,j} \sqrt{p_i p_j} |\phi_i^R\rangle\langle\phi_j^R| \otimes \hat{S}(|\psi_i\rangle\langle\psi_j|). \quad (21)$$

Where $S(\rho)$ is the entropy of von Newman and ψ^R is a purification of ρ

$$|\psi^R\rangle = \sum_i \sqrt{p_i} |\psi_i\rangle \otimes |\phi_i^R\rangle, \quad \langle\phi_j^R|\phi_i^R\rangle = \delta_{ij}, \quad (22)$$

$$\text{tr}_R |\psi^R\rangle\langle\psi^R| = \rho, \quad (23)$$

here $\{|\phi_i^R\rangle\}$ is some orthonormal set. The definition is independent from concrete choice of this set [6]. The mutual quantum information is the decrease of the entropy after acting of \hat{S} due to the possible distortion of entanglement state. This quantity is not symmetric with respect to interchanging of input and output and can be positive, negative or zero in contrast with the Shannon mutual information in classical theory.

It has been shown that (20) can be the upper bound of the capacity of a quantum channel [8,18]. Using this value the authors [8] have been proved the converse coding theorem for quantum source with respect to the entanglement fidelity [6]. Only this fidelity is adequate for quantum data transmission or compression. In the [7] the authors prove data processing inequality

$$I(\rho; \hat{S}_1) \geq I(\rho; \hat{S}_2 \hat{S}_1). \quad (24)$$

In [8] we found the alternative derivation of this result which is more simple than derivation of [7]. In this paper we show that this equation can be strengthened. Data processing inequality is very important property of mutual information. This is an effective tool for proving general results and the first step toward identification a physical quantity as mutual information.

Now we brief recall the derivation of data processing inequality in the general case. The formalism of relative quantum entropy is very useful in this context [11,13].

We have

$$\begin{aligned} & S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|) || \hat{1}^R \otimes \hat{S}(\rho^R \otimes \rho)) \\ &= -S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|)) + S(\rho^R) + S(\hat{S}\rho). \end{aligned} \quad (25)$$

Here

$$\rho^R = \sum_{i,j} \sqrt{p_i p_j} |\phi_i^R\rangle\langle\phi_j^R| \langle\psi_i|\psi_j\rangle. \quad (26)$$

Now from Lindblad inequality we have

$$\begin{aligned} & S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|) || \hat{1}^R \otimes \hat{S}(\rho^R \otimes \rho)) \\ & \geq S(\hat{1}^R \otimes \hat{S}_1 \hat{S}_2(|\psi^R\rangle\langle\psi^R|) || \hat{1}^R \otimes \hat{S}_1 \hat{S}_2(\rho^R \otimes \rho)). \end{aligned} \quad (27)$$

From this formula we have (24). From (25) we have

$$I(\rho; c\hat{S}_1 + (1-c)\hat{S}_2) \leq cI(\rho; \hat{S}_1) + (1-c)I(\rho; \hat{S}_2)$$

This theorem have been proved in [18].

Now we can prove the strengthened data processing inequality. Let in (27) \hat{S}_2 is represented in the form (11). From (6,11) we get

$$\begin{aligned} & S(\hat{1}^R \otimes \hat{S}_2 \hat{S}_1(|\psi^R\rangle\langle\psi^R|) || \hat{1}^R \otimes \hat{S}_2 \hat{S}_1(\rho^R \otimes \rho)) \\ & \leq -(1-c)S(\hat{1}^R \otimes \hat{C}_2 \hat{S}_1(|\psi^R\rangle\langle\psi^R|)) + S(\rho^R) + (1-c)S(\hat{C}_2 \hat{S}\rho). \end{aligned} \quad (28)$$

And we have

$$(1 - c(\hat{S}_2))I(\rho; \hat{S}_1) \geq I(\rho; \hat{S}_2 \hat{S}_1). \quad (29)$$

Now we consider the simplest example of noisy quantum channel: Two dimensional, two- Pauli channel [15]

$$A_1 = \sqrt{x}\hat{1}, \quad A_2 = \sqrt{(1-x)/2}\sigma_1, \quad A_3 = -i\sqrt{(1-x)/2}\sigma_2, \quad 0 \leq x \leq 1, \quad (30)$$

where $\hat{1}$, σ_1 , σ_2 are the unit matrix and the first and the second Pauli matrices. The (30) has also physical meaning as evolution operator for two-dimensional open system.

Any density matrix in two-dimensional Hilbert space can be represented in the Bloch form

$$\rho = (1 + \mathbf{a}\boldsymbol{\sigma})/2, \quad (31)$$

where \mathbf{a} is a real vector with $|\mathbf{a}| \leq 1$. Now we have

$$\hat{S}_{TP}((1 + \mathbf{a}\boldsymbol{\sigma})/2) = (1 + \mathbf{b}\boldsymbol{\sigma})/2, \quad (32)$$

where $\mathbf{b} = (a_1x, a_2x, a_3(2x - 1))$. After simple calculations we get

$$c(\hat{S}_{TP}) = (1 - |2x - 1|)/2. \quad (33)$$

We conclude by reiterating the main results: Lindblad inequality can be generalized. We have results not only about increasing of the entropy and decreasing of the mutual quantum information but also about velocity of these processes.

References

1. F.Shlogl, Phys. Rep., **62**, 268, (1980).
2. R.L. Stratanovich, *Information Theory*, 1975, Moscow, Nauka.
3. R.L. Stratanovich, *Nonequilibrium, Nonlinear Thermodynamics*, 1985, Moscow, Nauka.
4. R.P. Poplavsky, *Thermodynamics of information processes*, 1981, Moscow, Nauka.
5. S.M. Barnett and S.J.D. Phoenix, Phys. Rev. A **44**, 535, (1991).
6. B.Schumacher, Phys. Rev. A **54**, 2614, (1996).
7. B.Schumacher and M.A. Nielsen, Phys. Rev. A **54**, 2629, (1996).
8. A.E. Allahverdyan and D.B. Saakian, eprint quant-ph/9702023.
9. S. Lloyd, Phys. Rev. A **55**, R1613-1622, (1997); also e-print quant-ph/9604015.
10. K.Kraus, Ann. Phys., **64**, 311, (1971).
11. M.Ohya, Rep. Math. Phys., **27**, 19, (1989).
12. H.Umegaki, Kodai Math. Sem. Rep., **14**, 59, (1962).
13. G.Lindblad, Commun. Math. Phys., **40**, 147, (1975).
14. S.J.D. Phoenix and P.L. Knight, Ann. Phys.(N.Y) **186**, 381, (1988).
15. C.H. Bennett, C.A. Fuchs, J.A. Smolin, eprint quant-ph/9611006.
16. I.Csiszar, J.Korner, *Information Theory*, 1982.
17. F.Gantmacher, *The Theory of matrices*, Moscow, Nauka 1983.
18. H.Barnum, M.P. Nielsen, B.Schumacher, eprint quant-ph/9702049.

Fault-Tolerant Quantum Computation with Higher-Dimensional Systems

Daniel Gottesman

T-6 Group, Los Alamos National Laboratory
gottesma@t6-serv.lanl.gov

Abstract. Instead of a quantum computer where the fundamental units are 2-dimensional qubits, we can consider a quantum computer made up of d -dimensional systems. There is a straightforward generalization of the class of stabilizer codes to d -dimensional systems, and I will discuss the theory of fault-tolerant computation using such codes. I prove that universal fault-tolerant computation is possible with any higher-dimensional stabilizer code for prime d .

1 Introduction

Quantum computation and quantum communications have the potential to accomplish many things that would be difficult or impossible using just classical computers and communications. However, quantum data is very vulnerable to decoherence and to errors. It is likely that some form of quantum error correction will be needed to perform anything beyond the simplest computations with a quantum computer. Quantum error-correcting codes [1,2,3,4,5] provide one of the tools necessary. Such a code can protect quantum data against errors occurring during transmission or storage of the data. However, to have a reliable quantum computer, we also need for the computation to be performed in a fault-tolerant manner [6]. Fault-tolerant quantum computation requires a protocol that not only maps states of a quantum code to other states of a quantum code, but prevents errors from propagating out of control.

A large group of useful codes was introduced in [3] and [4]. These codes are essentially the quantum equivalent of classical linear codes, in that they can be easily described and encoded, and that it is easy to measure the error syndrome (as in the classical case, it may be difficult to compute the actual error from the error syndrome). The primary complication involved in quantum error correction is that it is not only necessary to correct bit flip errors

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (1)$$

but also phase errors

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2)$$

Consequently, it turns out to be useful to look at the group \mathcal{P} generated by tensor products of these operators. \mathcal{P} is called the Pauli group or the extraspecial group. *Stabilizer codes* are those codes where the valid codewords are all eigenstates of $n-k$ operators in \mathcal{P} . These $n-k$ operators generate a 2^{n-k} element Abelian group, called the *stabilizer* S of the code. The set of valid codewords forms a 2^k -dimensional subspace of the full n -qubit Hilbert space, the *coding space* of the code. The stabilizer is analogous to the parity check matrix of a classical linear code. In fact, the set of classical linear binary codes is exactly the set of stabilizer codes where everything in the stabilizer is a tensor product of Z operators.

Stabilizer codes are easy to work with because of the structure of the Pauli group.

$$XZ = -ZX \quad , \quad (3)$$

so any two operators in \mathcal{P} either commute or they anticommute. If an operator E anticommutes with an operator G in the stabilizer S of a code, then $E|\psi\rangle$ will have eigenvalue -1 for G instead of $+1$. Therefore, by measuring the eigenvalues of the $n-k$ generators of S , we can easily measure the error syndrome, and if the code is suitably chosen, identify the error.

Even if we restrict our attention to stabilizer codes, it is far from obvious that we can perform fault-tolerant computation. First of all, we must find some operators that map the coding space of the code into itself. Secondly, many of these operators will cause errors to spread from one qubit in a block to a different qubit in the same block of the code. Therefore, a single error could rapidly grow to become many errors within a block, exceeding the code's capability to correct them. In order to avoid this, we will further restrict attention to *transversal* operations, that is, operations which only interact qubits from one block with corresponding qubits in other blocks. This means that an error occurring in qubit number 3 in one block might spread to qubit number 3 in another block, but it will never spread back to qubit 2 in the first block. Each block can easily correct a single error, so this situation causes no problems. Shor was the first to demonstrate a universal set of fault-tolerant gates [6]. However, his construction only worked for a small class of codes. In [7], I was able to show that Shor's construction could be generalized to any qubit stabilizer code. The proof made extensive use of the group of unitary operators that leave the group \mathcal{P} invariant under conjugation. This group is known as the Clifford group.

In the classical theory of error-correcting codes, it is often helpful to go beyond bits and work with higher-dimensional systems. The same may be true for quantum error correction. Instead of a system made up of 2-dimensional qubits, we can work with a system composed of d -dimensional *qudits*. It turns out that there is a natural generalization of stabilizer codes to higher-dimensional systems [8,9,10]. In Sec. 2 I will present this generalization. Then I will proceed to generalize the arguments of [7] to show that universal fault-tolerant computation is also possible with any of these codes, at least in the case where d is prime. Though I will largely focus on prime d , I will also say a little bit about the general case along the way. Assume that d is prime unless it is otherwise specified.

The construction of a universal set of gates used in [7] consisted of a number of steps:

1. The full Clifford group can be constructed given the controlled-NOT, operators in the Pauli group, and measurement of operators in the Pauli group.
2. For any stabilizer code, we can perform encoded versions of all operators in the Pauli group and can measure operators in the Pauli group.
3. We can perform a CNOT between corresponding encoded qubits in different blocks of the code.
4. We can swap individual encoded qubits from one block of the code to an empty block and vice-versa.
5. We can move an encoded qubit in an otherwise empty block to whatever position in the block we desire. At this point, we have established that we can perform a CNOT between any pair of encoded qubits in the same or different blocks, and therefore can perform the full Clifford group.
6. Given the full Clifford group, we can perform an additional gate outside the Clifford group, such as the Toffoli gate or the $\pi/8$ rotation. This completes the universal set of gates.

Each of the steps in this construction has an analog for higher-dimensional systems. However, in this paper, I will present a simplified construction that combines steps [3] through [5]. This construction gives a direct way to perform the generalization of the CNOT between any pair of encoded qudits, whether they are in the same or different blocks and whether they are in corresponding or different positions within their blocks. This construction can also be used to simplify the proof for systems with $d = 2$, and will likely reduce the required overhead for fault-tolerant computation using codes with many qudits per block.

2 Higher Dimensional Generalization of the Pauli Group and Other Structures

The Pauli group has a natural generalization to higher-dimensional systems.^[1] Instead of generating it from the two-dimensional X and Z , we instead generate \mathcal{P} from tensor products of X_d and Z_d , where $X_d|j\rangle = |j+1\rangle$ and $Z_d|j\rangle = \omega^j|j\rangle$, where ω is a primitive d -th root of unity. X_d and Z_d satisfy the relation

$$X_d Z_d = \omega^{-1} Z_d X_d . \quad (4)$$

Both X_d and Z_d have order d . The elements of the single-qudit Pauli group have the form $\omega^a X_d^r Z_d^s$, where $0 \leq r, s < d$, and

$$(X_d^r Z_d^s) (X_d^t Z_d^u) = \omega^{st - ru} (X_d^t Z_d^u) (X_d^r Z_d^s) . \quad (5)$$

¹ The group presented in this section is not the only generalization of the Pauli group, although it is probably the simplest. See [8] for a more extensive discussion of this issue.

\mathcal{P} for n qubits will contain d^{2n} elements, plus an additional factor of d for overall phase. The elements of \mathcal{P} have eigenvalues ω^r for $r = 0, \dots, d-1$.² From now on, I will suppress the subscript d , and all operations should be taken to be over qudits instead of qubits.

The d -dimensional generalization of the stabilizer S of a code is again just an Abelian subgroup of \mathcal{P} . The coding space is composed of those states that are fixed by all elements of S (when d is even, this actually imposes an additional constraint on the overall phase of elements of S). If the stabilizer on n qudits has $n-k$ generators, then S will have d^{n-k} elements and the coding space will consist of k qudits. Note that this last fact need no longer be true when d is not prime, and this is the main source of complications in that case. It is unclear exactly how to deal with a code that does not encode an integral number of qudits. If we stick to codes for which all the generators of the stabilizer have order d , the rest of the proof will hold, modulo a question about gates necessary to generate the Clifford group.

If an operator E and $M \in S$ satisfy

$$EM = \omega^a ME, \quad (6)$$

then $E|\psi\rangle$ will have eigenvalue ω^a for M instead of eigenvalue $+1$, so we can detect that error E has occurred by measuring the eigenvalue of M .

We can see the structure of the coding space by extending the generators of S to a complete independent set of commuting operators. When d is not prime, we also require these operators to have order d . Such a set will have cardinality n , so we can do this by choosing k additional operators $\overline{Z}_1, \dots, \overline{Z}_k$. These operators have the interpretation of the encoded Z operators for the k encoded qudits. We can then choose k more operators $\overline{X}_1, \dots, \overline{X}_k$ which satisfy the relations

$$\overline{X}_i \overline{Z}_j = \overline{Z}_j \overline{X}_i \quad (i \neq j) \quad (7)$$

$$\overline{X}_i \overline{Z}_i = \omega^{-1} \overline{Z}_i \overline{X}_i \quad (8)$$

$$\overline{X}_i M = M \overline{X}_i \quad (\forall M \in \mathcal{P}) \quad (9)$$

The operators \overline{X}_i then act as the encoded X operators for the k encoded qudits. The generators of S along with $\overline{Z}_1, \dots, \overline{Z}_k$ and $\overline{X}_1, \dots, \overline{X}_k$ then generate the group of all Pauli group operators that commute with S . As in the two-dimensional case, the operators that commute with S but are not themselves in S are precisely the operators that cannot be detected by the quantum code. They therefore perform encoded operations on the data.

The Clifford group is the set of operators that leave \mathcal{P} invariant under conjugation. That is, it is the normalizer $N(\mathcal{P})$ of \mathcal{P} in the unitary group $U(d^n)$. The Clifford group is important for fault-tolerant computation because if we perform an operator U on the Hilbert space, the operator UNU^\dagger has the same relationships to states after the transformation as the operator N did before the

² This is true for odd d . For even d , XZ has order $2d$, so extra factors of i will be necessary, as in the $d = 2$ case. This aspect is actually simpler for odd d than for $d = 2$.

transformation. Therefore, instead of considering transformations of the states $|\psi\rangle \rightarrow U|\psi\rangle$, we can consider transformations of the operators $N \rightarrow UNU^\dagger$. When U is in the Clifford group and N is in the Pauli group, then UNU^\dagger will also be in the Pauli group. Therefore, we can uniquely describe elements of $N(\mathcal{P})$ by the permutation they induce on \mathcal{P} . The permutation must preserve the group structure of \mathcal{P} , but is otherwise arbitrary.

In the two-dimensional group, $N(\mathcal{P})$ was generated by two single-qubit operators R (the Hadamard transform $|j\rangle \rightarrow |0\rangle + (-1)^j|1\rangle$) and P (the phase gate $|j\rangle \rightarrow i^j|j\rangle$), and the two-qubit operator CNOT ($|i\rangle|j\rangle \rightarrow |i\rangle|(i+j) \bmod 2\rangle$). In d dimensions, R generalizes to the d -dimensional discrete Fourier transform

$$|j\rangle \rightarrow \sum_{s=0}^d \omega^{js} |s\rangle , \quad (10)$$

P generalizes to the d -dimensional phase gate

$$|j\rangle \rightarrow \omega^{j(j-1)/2} |j\rangle , \quad (11)$$

and CNOT generalizes to the SUM gate

$$|i\rangle|j\rangle \rightarrow |i\rangle|(i+j) \bmod d\rangle . \quad (12)$$

We can describe these operators by their induced transformations on the Pauli group. R maps

$$X \rightarrow Z , \quad (13)$$

$$Z \rightarrow X^{-1} . \quad (14)$$

P maps

$$X \rightarrow XZ , \quad (15)$$

$$Z \rightarrow Z . \quad (16)$$

SUM maps

$$X \otimes I \rightarrow X \otimes X , \quad (17)$$

$$I \otimes X \rightarrow I \otimes X , \quad (18)$$

$$Z \otimes I \rightarrow Z \otimes I , \quad (19)$$

$$I \otimes Z \rightarrow Z^{-1} \otimes Z . \quad (20)$$

However, it is not clear that these three gates generate the Clifford group. We may also need the S gate

$$X \rightarrow X^a , \quad (21)$$

$$Z \rightarrow Z^b , \quad (22)$$

for all pairs (a, b) , where $ab \equiv 1 \bmod d$. On kets, this gate acts as $|j\rangle \rightarrow |aj\rangle$. In fact, a single pair (a, b) is sufficient, as long as a generates the multiplicative

group \mathbb{Z}_d^* . I will not give a detailed proof that these gates generate the Clifford group, but using the P , R , and S gates, we can get the full one-qudit Clifford group. Then a construction similar to that used in [7] will give the full n -qudit Clifford group. The structure is somewhat more complicated when d is not prime, and I have not verified that these gates are sufficient for the nonprime case.

Note that we can fault-tolerantly measure any operator in \mathcal{P} that is the tensor product of Z operators by performing transversal SUM gates from the qudits to be measured to an appropriate ancilla. Since we are interested in eigenvalues with possible values ω^j for $j = 0, \dots, d-1$, the appropriate ancilla state is the superposition of all states where the sum of the qudits is $0 \bmod d$. That way, no information beyond the eigenvalue of the measured operator will be conveyed. We can construct this state by Fourier transforming the state $\sum_{j=0}^{d-1} |jj \cdots j\rangle$. Following DiVincenzo and Shor [11], we can measure any operator in \mathcal{P} by performing a transversal Clifford group operation C that takes the operator to the tensor product of Z 's, performing the measurement, and applying C^{-1} .

For any stabilizer code, the elements of the k -qudit encoded Pauli group are also elements of the n -qudit unencoded Pauli group, as are the generators of the stabilizer. Since we can perform and measure an arbitrary element of the unencoded Pauli group, we have shown that for a stabilizer code over d -dimensions, we can apply encoded versions of X and Z for all encoded qudits, measure the generators of the stabilizer (and therefore perform fault-tolerant error correction), and measure all members of the encoded Pauli group fault-tolerantly. This provides step 2 of the proof.

3 Measurements and Stabilizers

In [7], it proved very helpful in a number of places to understand how the stabilizer of a state or subspace changed under measurements. The procedure for qubits generalizes easily to higher dimensions.

First, recall that there is more than one way to choose generators for a given stabilizer. Any maximal set of independent operators in the group will suffice. In particular, any generator M can be replaced by NM for any $N \neq M$. Similarly, the encoded \overline{X} and \overline{Z} operators are only defined up to multiplication by elements of S . If we wish to measure an operator $A \in \mathcal{P}$, then the first step is to put the stabilizer and \overline{X} and \overline{Z} operators in a form so that all the \overline{X} and \overline{Z} operators commute with A and all but one of the generators of S commutes with A . We can do this because if $M \in S$ does not commute with A , then $M^a N$ will commute with A for some a for any N that commutes with M (as do all \overline{X} 's, \overline{Z} 's, and generators of S). We will not need to consider the case where A commutes with everything in the stabilizer.

This is a useful form for the stabilizer because any operator that commutes with A is not disturbed by the measurement of A . Therefore, we only need to change M when A is measured. Since $A \in \mathcal{P}$, the possible measurement results are ω^a for $a = 0, \dots, d-1$. These result ω^a corresponds to applying the projection operator

$$P_a = \left(I + \omega^{-a} A + \omega^{-2a} A^2 + \dots + \omega^{-(d-1)a} A^{d-1} \right) / d \quad (23)$$

to the state. Assume now that

$$MA = \omega AM \quad (24)$$

(note that when d is prime, this will always be true for some power of M). Then

$$MP_a M^\dagger = \frac{1}{d} \sum_{j=0}^{d-1} \omega^{-ja} M A^j M^\dagger = \frac{1}{d} \sum_{j=0}^{d-1} \omega^{-j(a-1)} A^j = P_{a-1} . \quad (25)$$

Thus, if we measure A and get the result a , by applying M^a we can produce the same state we would have gotten if we had gotten the result 0. I will assume below that any measurement is followed by such a correction. Once this correction is performed, the new state is a $+1$ eigenvector of A , so A should be added to the stabilizer. It is not an eigenvector of M , so M is removed from the stabilizer. All of the other generators (which have been put in a form where they commute with A) are unchanged.

4 Gates Derived from SUM

Suppose we have the ability to perform the SUM gate between any pair of qudits in our computer, as well as the ability to perform the Pauli group and to measure operators in the Pauli group. I will now show that we can apply the full Clifford group to the computer.

Suppose we consider a single unknown qudit and prepare a second ancilla qudit in the state $|0\rangle$. This two-qudit system can be described by the stabilizer $I \otimes Z$. The logical Pauli group is generated by $\overline{X} = X \otimes I$ and $\overline{Z} = Z \otimes I$. Now perform a SUM gate from the first qudit to the second qudit. The stabilizer is $Z^{-1} \otimes Z$, $\overline{X} = X \otimes X$, and $\overline{Z} = Z \otimes I$.

Suppose we were now to measure the operator $A = I \otimes XZ$. Then $M = Z^{-1} \otimes Z \in S$ and $MA = \omega AM$. Therefore, this measurement results in the stabilizer $I \otimes XZ$, and $\overline{X} = XZ^{-1} \otimes XZ$ and $\overline{Z} = Z \otimes I$. We can discard the second qudit, and the effective transformation on the first qudit is

$$X \rightarrow XZ^{-1} , \quad (26)$$

$$Z \rightarrow Z . \quad (27)$$

This is the gate P^{-1} . $d-1$ iterations of it will produce the P gate.

Alternatively, we could have prepared the ancilla qudit so that the stabilizer of the system began as $I \otimes X$, then performed the SUM gate from the second qudit to the first. The stabilizer would then be $X \otimes X$, and $\overline{X} = X \otimes I$ and $\overline{Z} = Z \otimes Z^{-1}$. Then we measure $A = I \otimes XZ^{-1}$ and choose $M = X \otimes X$ so that $MA = \omega AM$. The final stabilizer is $I \otimes XZ^{-1}$, so we discard the second qudit, leaving

$$X \rightarrow X , \quad (28)$$

$$Z \rightarrow XZ . \quad (29)$$

Call this gate Q . Then $R^{-1} = XQP^{-1}Q$, and $R = R^{-3}$.

Now suppose we again prepare the ancilla in the $+1$ eigenstate of X , but now perform s SUM gates from the second qudit to the first instead of one. The stabilizer is $X^s \otimes X$, $\overline{X} = X \otimes I$, and $\overline{Z} = Z \otimes Z^{-s}$. This time we measure $A = Z \otimes I$. This results in stabilizer $Z \otimes I$,

$$\overline{X} = (X \otimes I)(X^s \otimes X)^{-s^{-1}} = I \otimes X^{-s^{-1}}, \quad (30)$$

and $\overline{Z} = Z \otimes Z^{-s}$. Therefore, discarding the first qudit leaves the transformation

$$X \rightarrow X^{-s^{-1}}, \quad (31)$$

$$Z \rightarrow Z^{-s}. \quad (32)$$

By choosing an appropriate s , we can therefore perform an arbitrary S gate. Note that in this case, the data ends up in what was originally the ancilla qudit.

I have shown how to produce the P , R , and S gates from the SUM gate. Therefore, given the SUM gate, we can produce the full Clifford group. This completes step [II](#) of the proof.

5 Producing the SUM Gate for Any Stabilizer Code

To see how to construct the SUM gate between any pair of encoded qudits, first consider two unencoded qudits. Introduce a third qudit in the state $|0\rangle$. The stabilizer at this point is $I \otimes I \otimes Z$. Assume we can do Pauli group measurements, even entangled ones, and perform operators in the Pauli group. Let us first measure the operator $I \otimes X \otimes X^{-1}$. This becomes the stabilizer. The logical Pauli group generators are

$$\overline{X}_1 = X \otimes I \otimes I, \quad (33)$$

$$\overline{X}_2 = I \otimes X \otimes I, \quad (34)$$

$$\overline{Z}_1 = Z \otimes I \otimes I, \quad (35)$$

$$\overline{Z}_2 = I \otimes Z \otimes Z. \quad (36)$$

Now measure $Z \otimes I \otimes Z$. It becomes the new stabilizer, and

$$\overline{X}_1 = X \otimes X \otimes X^{-1}, \quad (37)$$

$$\overline{X}_2 = I \otimes X \otimes I, \quad (38)$$

$$\overline{Z}_1 = Z \otimes I \otimes I, \quad (39)$$

$$\overline{Z}_2 = I \otimes Z \otimes Z. \quad (40)$$

Finally, measure $I \otimes I \otimes X$ and discard the last qudit. This leaves us with

$$\overline{X}_1 = X \otimes X, \quad (41)$$

$$\overline{X}_2 = I \otimes X, \quad (42)$$

$$\overline{Z}_1 = Z \otimes I, \quad (43)$$

$$\overline{Z}_2 = Z^{-1} \otimes Z. \quad (44)$$

This we recognize as the transformation induced by the SUM gate, so this series of entangled measurements has performed the SUM gate between these two qudits.

Now, to apply this to a quantum code, we just need to be able to measure entangled logical Pauli group operators between any pair of encoded qudits. If the qudits are in the same block, this is straightforward. For instance, if they are in slots i and j , the encoded version of $X \otimes X$ is just $\bar{X}_i \bar{X}_j$. This is in the Pauli group too, so we know how to measure it.

If the qudits are in different blocks, it is not much harder. Instead of using an a -qudit ancilla state, we use an $(a+b)$ -qudit ancilla state (where a and b are the weights of the operators \bar{X}_i and \bar{X}_j), which is again in the superposition of all states whose registers sum to $0 \bmod d$. The operator we wish to measure is $\bar{X}_i \otimes \bar{X}_j$, which is in the Pauli group. By performing the appropriate transversal Clifford group operation, we rotate this to be the tensor product of Z 's and perform SUM gates from the appropriate qudits to the corresponding ancilla qudits, then perform the inverse Clifford group operator to restore the state to its original form. Then we measure the $a+b$ ancilla qudits, and this tells us the eigenvalue of the measured operator. We use an $(a+b)$ -qudit ancilla instead of an a -qudit ancilla plus a b -qudit ancilla because we do not wish to be able to find the eigenvalues of $\bar{X}_i \otimes I$ and $I \otimes \bar{X}_j$ separately, only their product.

Therefore, given an encoded ancilla qudit which is initialized to $|0\rangle$, by performing the encoded version of the above entangled measurements, we can perform a SUM gate between any pair of encoded qudits anywhere in the computer. Note that the ancilla qudit can itself be anywhere in the computer; it need not be in the same block as either data qudit, or in the corresponding place in a different block.

Given the SUM gate and the results of the previous section, we can perform the full Clifford group on the encoded data for any stabilizer code. This completes the proof up to step [5](#). This part of the proof is a significant improvement on the method used in [7](#). In that paper, it was necessary to introduce full ancilla blocks to perform a CNOT. Here, we need only a single logical ancilla qudit. In the case where a block may encode many qudits, this can be a major improvement. The price is that we must potentially perform entangled measurements on more than one block. This means we will have to use larger ancilla states for the measurement; this results in a greater potential for error, so we will have to repeat the measurement more times, and perhaps perform error correction a bit more often. However, in many situations, the total number of physical gates we need will decrease.

Note that this procedure works just as well if the two logical qudits involved in the SUM are in blocks made up of different numbers of physical qudits. This means we can interact qudits encoded with different sorts of codes fault-tolerantly, or change the encoding of a single qudit without losing the protection against errors at any time.

6 Completing the Universal Set of Gates

The set of universal gates can be completed by adding the higher-dimensional analog of the Toffoli gate [12]

$$|a\rangle|b\rangle|c\rangle \rightarrow |a\rangle|b\rangle|c+ab\rangle . \quad (45)$$

It turns out that a generalization of Shor's fault-tolerant construction of the Toffoli gate [6] will work here.

Suppose we prepare a three-qudit ancilla in the $+1$ eigenstate of the three operators

$$M_1 = (X \otimes I \otimes I) \text{SUM}(2 \rightarrow 3) , \quad (46)$$

$$M_2 = (I \otimes X \otimes I) \text{SUM}(1 \rightarrow 3) , \quad (47)$$

$$M_3 = (I \otimes I \otimes Z) \text{PHASE}(1, 2)^{-1} . \quad (48)$$

$\text{SUM}(i \rightarrow j)$ is a SUM gate performed with the i th qudit as control and the j th qudit as target. $\text{PHASE}(i, j)$ is the PHASE gate

$$\text{PHASE}|a\rangle|b\rangle = \omega^{ab}|a\rangle|b\rangle \quad (49)$$

performed on the i th and j th qudits. One important fact to note is that both of these gates are in the Clifford group. Since we have already constructed the Clifford group, this will enable us to also construct the Toffoli gate. The appropriate state is

$$|A\rangle = \sum_{a,b} |a\rangle|b\rangle|ab\rangle . \quad (50)$$

Now, given three data qudits, perform inverse SUM gates (i.e., $|a\rangle|b\rangle \rightarrow |a\rangle|b-a\rangle$) from the first and second ancilla qudits to the first and second data qudits, respectively, and a SUM gate from the third data qudit to the third ancilla qudit. Now we measure the last three qudits, the original data qudits, in the bases Z , Z , and X , respectively. After performing the appropriate corrections (which will likely involve gates from the Clifford group as well as the Pauli group), we are left with the data in the first three qudits, which were originally the ancilla qudits. It turns out that after these operations, a Toffoli gate has been performed on the data qudits.

To construct the appropriate ancilla state $|A\rangle$, we can again follow Shor. The various states

$$|A_j\rangle = \sum_{a,b} |a\rangle|b\rangle|ab+j\rangle \quad (51)$$

for $j = 0, \dots, d-1$ are related to $|A\rangle$ by

$$|A_j\rangle = (I \otimes I \otimes X^j) |A\rangle . \quad (52)$$

The states $|A_j\rangle$, like $|A\rangle$, are $+1$ eigenstates of M_1 and M_2 , but $M_3|A_j\rangle = \omega^j|A_j\rangle$.

Furthermore, note that

$$\sum_{j=0}^{d-1} |A_j\rangle = \sum_{a,b,c} |a\rangle|b\rangle|c\rangle . \quad (53)$$

This last state is easily constructed as the Fourier transform of $|000\rangle$. Then by measuring M_3 for this state, we can collapse the state into one of the states $|A_j\rangle$. By applying the operator X^{-j} , we get the state $|A\rangle$.

To measure M_3 fault-tolerantly for a quantum code, we prepare the CAT state $\sum_j |jj \cdots j\rangle$ (using the same number of qudits as in a block of the code). M_3 is in the Clifford group, so we can perform it by some sequence of transversal operations and measurements. By conditioning the appropriate operations for the i th qudit on the i th qudit of the CAT state, we can conditionally perform M_3 on the code depending on the CAT state. Note that a conditional operation in the d -dimensional case means applying M_3^j when the control qudit is in the state $|j\rangle$. Once we have done this, when the code is in an eigenstate of M_3 with eigenvalue ω^s , the CAT state ends up in the state

$$|CAT_s\rangle = \sum_j \omega^{js} |jj \cdots j\rangle . \quad (54)$$

The various states $|CAT_s\rangle$ are orthogonal to each other, and so can be distinguished by an appropriate measurement. This therefore gives us a measurement of M_3 , completing the construction of the Toffoli gate and step 6 of the proof.

Because everything we do is transversal, single qudit errors in the CAT state cannot become more than single qudit errors in any single block of the code. Naturally, after creating the CAT state we should verify it to make sure there are no correlated errors. In addition, a single qudit error in the CAT state could give us the wrong measurement result. Therefore, the measurement of M_3 should be repeated in order to sufficiently increase our confidence in the result.

Acknowledgements

I would like to thank Manny Knill and Raymond Laflamme for helpful discussions, and Michael Nielsen for suggesting the name ‘‘Pauli group.’’

References

1. P. Shor, Phys. Rev. A **52**, 2493 (1995).
2. A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).
3. D. Gottesman, Phys. Rev. A **54**, 1862 (1996).
4. A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Phys. Rev. Lett. **78**, 405 (1997).
5. A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, ‘‘Quantum error correction via codes over $\text{GF}(4)$,’’ quant-ph/9608006, to appear in IEEE Trans. Information Theory.

6. P. Shor, *Proceedings of the 37th Symposium on the Foundations of Computer Science*, IEEE Computer Society Press (Los Alamitos, CA), 56 (1996).
7. D. Gottesman, Phys. Rev. A **57**, 127 (1998).
8. E. Knill, “Non-binary error bases and quantum codes,” quant-ph/9608048.
9. E. Knill, “Group representations, error bases and quantum codes,” quant-ph/9608049.
10. E. Rains, “Nonbinary quantum codes,” quant-ph/9703048.
11. D. DiVincenzo and P. Shor, Phys. Rev. Lett. **77**, 3260 (1996).
12. E. Knill and R. Laflamme, private communication.

Quantum Convolutional Error Correction Codes

H.F. Chau

Department of Physics, University of Hong Kong, Pokfulam Road, Hong Kong
hfchau@hkusua.hku.hk

Abstract. I report two general methods to construct quantum convolutional codes for quantum registers with internal N states. Using one of these methods, I construct a quantum convolutional code of rate $1/4$ which is able to correct one general quantum error for every eight consecutive quantum registers. Keywords: Code Pasting, Convolutional Codes,

Phase Shift Error, Spin Flip Error, Quantum Codes, Quantum Error Correction.

1 Introduction

Quantum error correction code (QECC) is a succinct way to protect a quantum state from decoherence. The basic idea behind all QECC schemes is that by suitably encoding a quantum state in a larger Hilbert space H , and then later on measuring the wave function into certain subspace C of H , it is possible to detect the kind of errors that have occurred. Finally, one can correct the error by applying a suitable unitary transformation to the orthogonal complement of C according to the measurement result [23]. Many QECCs have been discovered in the last few years (see, for example, Refs. [4,6,7,8,9,10,15,18,19,22,23,24,25,26]) and various theories on the QECC have also been developed (see, for example, Refs. [3,4,7,9,10,15,16,17,19,21,22,25]). In particular, the necessary and sufficient condition for a QECC is [3,16,17]

$$\langle i_{\text{encode}} | \mathcal{A}^\dagger \mathcal{B} | j_{\text{encode}} \rangle = \Lambda_{\mathcal{A}, \mathcal{B}} \delta_{ij} , \quad (1)$$

where $|i_{\text{encode}}\rangle$ denotes the encoded quantum state $|i\rangle$ using the QECC; \mathcal{A}, \mathcal{B} are the possible errors the QECC can handle; and $\Lambda_{\mathcal{A}, \mathcal{B}}$ is a complex constant independent of $|i_{\text{encode}}\rangle$ and $|j_{\text{encode}}\rangle$.

All QECCs discovered so far are block codes. That is, the original state ket is first divided into *finite* blocks of the same length. Each block is then encoded separately using a code which is *independent* of the state of the other blocks (*cf.* Refs. [13,20]). Besides block codes, convolutional codes are well known in classical error correction. Unlike a block code, the encoding operation depends on current as well as a number of past information bits [13,20]. For instance, given a (possibly infinite) sequence of classical binary numbers $(a_1, a_2, \dots, a_m, \dots)$, the encoding $(b_1, c_1, b_2, c_2, \dots, b_m, c_m, \dots)$ with

$$b_i = a_i + a_{i-2} \bmod 2, \quad c_i = a_i + a_{i-1} + a_{i-2} \bmod 2 \quad (2)$$

for all i , and $a_0 = a_{-1} = 0$ is an example of classical convolutional code that can correct up to one error for every four consecutive bits (see, for example, chap. 4 in Ref. [13] and Lemma 3 in Section 3 for details).

In classical error correction, good convolutional codes often outperforms their corresponding block codes in the sense that they have higher encoding efficiencies [13, 20]. Thus, it is instructive to find quantum convolutional codes (QCC) and to analyze their performance. Here, I report two ways to construct QCCs. And from one of these methods, I construct a QCC of rate $1/4$ that can correct one quantum error for every eight consecutive quantum registers (see Ref. [11] for more details).

2 Constructing Quantum Convolutional Codes from Quantum Block Codes

In this Section, I report a general scheme to construct QCCs from quantum block codes (QBCs). But before doing so, let me first introduce some basic notations. Suppose each quantum register has N orthogonal eigenstates, where N is an integer greater than one. Then, the basis of a general quantum state making up of a collection of possibly infinite quantum registers can be chosen as $\{|\mathbf{k}\rangle\} \equiv \{|k_1, k_2, \dots, k_m, \dots\rangle\}$, where $k_m \in \mathbb{Z}_N$ for all $m \in \mathbb{Z}$ with $N \geq 2$. Moreover, I abuse the notation by defining $k_m = 0$ for all $m \leq 0$. Finally, all additions and multiplications in all state kets below are modulo N .

Definition 1. Let $|\mathbf{x}\rangle \equiv \sum_{k_1, k_2, \dots} a_{k_1, k_2, \dots} |k_1, k_2, \dots, k_m, \dots\rangle \equiv \sum_{\{\mathbf{k}\}} a_{\mathbf{k}} |\mathbf{k}\rangle$ be a quantum state. Any quantum error can be regarded as an error operator \mathcal{E} acting on this state. In particular, there is a **spin flip error** occurring at quantum register m (with respected to the basis $\{|\mathbf{k}\rangle\}$) if and only if $\mathcal{E}|\mathbf{x}\rangle = \sum_{\{\mathbf{k}\}} a_{\mathbf{k}} |k_1, k_2, \dots, k_{m-1}, \tilde{k}_m, k_{m+1}, \dots\rangle$, where $\tilde{k}_m(k_m, \mathcal{E})$ is a \mathbb{Z}_N -function of k_m and \mathcal{E} . Moreover, a spin flip error is said to be **additive** provided that $\tilde{k}_m(k_m, \mathcal{E}) = k_m + \alpha \bmod N$ for some $\alpha(\mathcal{E})$.

Similarly, there is a **phase shift error** occurring at quantum register m (with respected to the basis $\{|\mathbf{k}\rangle\}$) if and only if $\mathcal{E}|\mathbf{x}\rangle = \sum_{\{\mathbf{k}\}} a_{\mathbf{k}} f(k_m, \mathcal{E}) |\mathbf{k}\rangle$ for some complex-valued function $f(k_m, \mathcal{E})$ with $|f|^2 = 1$. Spin flip and phase shift errors occurring at more than one quantum register are defined in a similar way.

With the above notations and definition in mind, a QBC and a QCC can be defined as follows:

Definition 2. The linear map sending

$$\begin{aligned} |\mathbf{k}\rangle &\equiv |k_1, k_2, \dots, k_n\rangle \\ \longmapsto \sum_{i_1, i_2, \dots, i_m} a_{i_1, i_2, \dots, i_m}^{(\mathbf{k})} |i_1, i_2, \dots, i_m\rangle &\equiv \sum_{\{\mathbf{i}\}} a_{\mathbf{i}}^{(\mathbf{k})} |\mathbf{i}\rangle \equiv |\mathbf{k}_{\text{encode}}\rangle, \end{aligned} \quad (3)$$

where $a_{\mathbf{i}}^{(\mathbf{k})} \in \mathbb{C}$, and $k_i \in \mathbb{Z}_N$ for all $i = 1, 2, \dots, N$ is said to be a **quantum block code (QBC)** that can correct errors in the set E if and only if Eq. (1)

is satisfied for all $\mathcal{A}, \mathcal{B} \in E$. Since Eq. (3) encodes ever n quantum registers to m registers, the **rate** of this code is, therefore, defined as n/m . In addition, one can encode the quantum state $\bigotimes_p |\mathbf{k}^{(p)}\rangle$ using the above QBC as $\bigotimes_p |\mathbf{k}_{\text{encode}}^{(p)}\rangle$.

On the other hand, if the encoding scheme expressed in Eq. (3) depends on current as well as past quantum states (that is, the coefficients $a_i^{(k)}$ in Eq. (3) depend on more than one $\mathbf{k}^{(p)}$), then it is called a **quantum convolutional code (QCC)**. The rate of this convolutional code equals n/m because it asymptotically encodes every n quantum registers as m registers.

With the above definitions in mind, one can construct a family of QCCs from a QBC as follows:

Theorem 1. Given a QBC in Eq. (3) and a quantum state $|\mathbf{k}\rangle \equiv \bigotimes_{i=1}^{+\infty} |\mathbf{k}_i\rangle$ making up of possibly infinitely many quantum registers, then the mapping

$$|\mathbf{k}\rangle \equiv \bigotimes_{i=1}^{+\infty} |\mathbf{k}_i\rangle \mapsto |\mathbf{k}_{\text{encode}}\rangle \equiv \bigotimes_{i=1}^{+\infty} \left[\sum_{\{j_i\}} a_{j_i}^{(\sum_p \mu_{ip} \mathbf{k}_p)} |j_i\rangle \right], \quad (4)$$

forms a QCC of rate n/m provided that the matrix μ_{ip} is invertible. This QCC handles errors in the set $E \otimes E \otimes E \otimes \dots$.

Proof. Let me consider the effects of errors $\mathcal{E} \equiv \mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \mathcal{E}_3 \otimes \dots$ and $\mathcal{E}' \equiv \mathcal{E}'_1 \otimes \mathcal{E}'_2 \otimes \mathcal{E}'_3 \otimes \dots$ in $E \otimes E \otimes E \otimes \dots$ on the encoded quantum registers by computing $\langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle$. From Eq. (1), I find that

$$\begin{aligned} \langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle &= \prod_{i=1}^{+\infty} \left[\sum_{\{j_i, j'_i\}} \bar{a}_{j'_i}^{(\sum_{p'} \mu_{ip'} \mathbf{k}'_{p'})} a_{j_i}^{(\sum_p \mu_{ip} \mathbf{k}_p)} \langle j'_i | \mathcal{E}_i'^\dagger \mathcal{E}_i | j_i \rangle \right] \\ &= \prod_{i=1}^{+\infty} \left[\left\langle \left(\sum_p \mu_{ip} \mathbf{k}'_p \right)_{\text{encode}} \left| \mathcal{E}_i'^\dagger \mathcal{E}_i \right| \left(\sum_p \mu_{ip} \mathbf{k}_p \right)_{\text{encode}} \right\rangle \right] \\ &= \prod_{i=1}^{+\infty} \left[\delta_{\sum_p \mu_{ip} \mathbf{k}_p, \sum_p \mu_{ip} \mathbf{k}'_p} \Lambda_{\mathcal{E}_i, \mathcal{E}'_i} \right] \end{aligned} \quad (5)$$

for some constants $\Lambda_{\mathcal{E}_i, \mathcal{E}'_i}$ independent of \mathbf{k} and \mathbf{k}' . Because μ is invertible, it is clear that $\mathbf{k}_i = \mathbf{k}'_i$ for all $i \in \mathbb{Z}^+$ is the unique solution for the systems of linear equations $\sum_p \mu_{ip} \mathbf{k}_p = \sum_p \mu_{ip} \mathbf{k}'_p$. Consequently, $\langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle = \delta_{\mathbf{k}, \mathbf{k}'} \Lambda_{\mathcal{E}, \mathcal{E}'}$, for some constant $\Lambda_{\mathcal{E}, \mathcal{E}'}$ independent of \mathbf{k} and \mathbf{k}' . Thus, the mapping in Eq. (4) is a QCC. \square

Now, let me uses Theorem 1 to give an example of QCC.

Example 1. Starting from the five qubit perfect code for $N = 2$ [6,10,18], Theorem 1 implies that the following QCC can correct up to one error in every five consecutive qubits:

$$|k_1, k_2, \dots, k_m, \dots\rangle \mapsto \bigotimes_{i=1}^{+\infty} \left[\frac{1}{N^{3/2}} \sum_{p_i, q_i, r_i=0}^{N-1} (-1)^{(k_i+k_{i-1})(p_i+q_i+r_i)+p_i r_i} |p_i, q_i, p_i + r_i, q_i + r_i, p_i + q_i + k_i + k_{i-1}\rangle \right] \quad (6)$$

where $k_m \in \{0, 1\}$ for all $m \in \mathbb{Z}^+$. The rate of this code is $1/5$.

Although the QCC in Eq. (3) looks rather complicated, the actual encoding process can be performed readily. Since μ is invertible, one can reversibly map $|k_1, k_2, \dots, k_n, \dots\rangle$ to $|\sum_p \mu_{1p} k_p, \sum_p \mu_{2p} k_p, \dots, \sum_p \mu_{np} k_p, \dots\rangle$ [12, 12]. Then, one obtains the above five bit QCC by encoding each quantum register using various encoding procedures described in Refs. [3, 5, 10, 18].

3 Constructing Quantum Convolutional Codes from Classical Convolutional Codes

In this Section, I report a general method to construct QCCs from classical convolutional codes. My construction is based on the following two technical lemmas which hold for both QBCs and QCCs:

Lemma 1. *Suppose the QECC*

$$|\mathbf{k}\rangle \mapsto \sum_{\{j\}} a_j^{(\mathbf{k})} |j\rangle \quad (7)$$

corrects (independent) additive spin flip errors in certain quantum registers. Then, the following QECC, which is obtained by discrete Fourier transforming every quantum register in Eq. (7),

$$|\mathbf{k}\rangle \mapsto \sum_{\{j, \mathbf{p}\}} a_j^{(\mathbf{k})} \prod_{i=1}^{+\infty} \left(\frac{1}{\sqrt{N}} \omega_N^{j_i p_i} \right) |\mathbf{p}\rangle \quad (8)$$

corrects (independent) phase errors occurring in the same set of quantum registers. The converse is also true.

Proof. Consider two arbitrary but fixed additive spin flip errors $\mathcal{E} \equiv \bigotimes_{i=1}^{+\infty} \mathcal{E}_i$ and $\mathcal{E}' \equiv \bigotimes_{i=1}^{+\infty} \mathcal{E}'_i$ acting on the code in Eq. (7). I denote the set of all quantum registers affected by either one of the above spin flip errors and unaffected by both errors as A and U , respectively. Then Eqs. (11) and (12) imply that

$$\sum_{\{j, j'\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \left(\prod_{i \in U} \delta_{j_i, j'_i} \right) \left(\prod_{i \in A} \langle j'_i | \mathcal{E}'_i^\dagger \mathcal{E}_i | j_i \rangle \right) \right] = \delta_{\mathbf{k}, \mathbf{k}'} \Lambda_{\mathcal{E}, \mathcal{E}'} \quad (9)$$

for some constant $\Lambda_{\mathcal{E}, \mathcal{E}'}$ independent of \mathbf{k} and \mathbf{k}' .

For additive spin errors, $\langle j'_i | \mathcal{E}'^\dagger \mathcal{E}_i | j_i \rangle = \langle j'_i + \alpha'_i | j_i + \alpha_i \rangle = \delta_{j'_i + \alpha'_i, j_i + \alpha_i}$ for some constants $\alpha_i, \alpha'_i \in \mathbb{Z}_N$. In other words, $\langle j'_i | \mathcal{E}'^\dagger \mathcal{E}_i | j_i \rangle$ is a binary function of $j_i - j'_i$ only. Thus, Eq. (9) still holds if I replace $\langle j'_i | \mathcal{E}'^\dagger \mathcal{E}_i | j_i \rangle$ by a binary function $g(j_i - j'_i : i \in A)$. Moreover, the linearity of Eq. (9) implies that the same equation holds if I replace $\langle j'_i | \mathcal{E}'^\dagger \mathcal{E}_i | j_i \rangle$ by *any* complex-valued function g taking arguments on $j_i - j'_i$ for all $i \in A$. That is to say,

$$\sum_{\{j, j'\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \left(\prod_{i \in U} \delta_{j_i, j'_i} \right) g(j_i - j'_i : i \in A) \right] = \delta_{\mathbf{k}, \mathbf{k}'} \Lambda_g \quad (10)$$

for some complex-valued Λ_g independent of \mathbf{k} and \mathbf{k}' . Conversely, it is obvious that if $a_j^{(\mathbf{k})}$ satisfies Eq. (10), then Eq. (7) is a QECC that is capable of correcting additive spin flip errors. In other words, Eq. (10) is a necessary and sufficient condition for the QECC to correct additive spin flip errors.

Now, I consider the actions of two phase shift errors \mathcal{F} and \mathcal{F}' acting on the *same* set of quantum registers as those in \mathcal{E} and \mathcal{E}' , respectively. Then

$$\begin{aligned} & \langle \mathbf{k}'_{\text{encode}} | \mathcal{F}'^\dagger \mathcal{F} | \mathbf{k}_{\text{encode}} \rangle \\ &= \sum_{\{j, j', \mathbf{p}, \mathbf{p}'\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \omega_N^{\sum_i (j_i p_i - j'_i p'_i)} \prod_{i \in U} \left(\frac{1}{\sqrt{N}} \delta_{p_i, p'_i} \right) \right. \\ & \quad \times \left. \prod_{i \in A} \left(\frac{1}{\sqrt{N}} \langle p'_i | \mathcal{F}'^\dagger_i \mathcal{F}_i | p_i \rangle \right) \right] \\ &= \sum_{\{j, j', \mathbf{p}\}} \left\{ \bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \left(\prod_{i \in U} \delta_{j_i, j'_i} \right) \sum_{\mathbf{p}' : i \in A} \left[\omega_N^{\sum_{i \in A} (j'_i p'_i - j_i p_i)} \right. \right. \\ & \quad \times \left. \left. \prod_{i \in A} \left(\frac{1}{\sqrt{N}} \langle p'_i | \mathcal{F}'^\dagger_i \mathcal{F}_i | p_i \rangle \right) \right] \right\}. \end{aligned} \quad (11)$$

For phase shift errors, $\langle p'_i | \mathcal{F}'^\dagger_i \mathcal{F}_i | p_i \rangle = \delta_{p_i, p'_i} h$ for some complex-valued function h of $p_i : i \in A$ with $|h|^2 = 1$. Consequently, Eq. (11) can be further simplified as

$$\langle \mathbf{k}'_{\text{encode}} | \mathcal{F}'^\dagger \mathcal{F} | \mathbf{k}_{\text{encode}} \rangle = \sum_{\{j, j', \mathbf{p}\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \left(\prod_{i \in U} \delta_{j_i, j'_i} \right) \omega_N^{\sum_{i \in A} p_i (j'_i - j_i)} h(p_i : i \in A) \right], \quad (12)$$

for some complex-valued function $h(p_i : i \in A)$. Summing over all the p_i s in Eq. (12), I obtain

$$\langle \mathbf{k}'_{\text{encode}} | \mathcal{F}'^\dagger \mathcal{F} | \mathbf{k}_{\text{encode}} \rangle = \sum_{\{j, j'\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \left(\prod_{i \in U} \delta_{j_i, j'_i} \right) h'(j_i - j'_i : i \in A) \right], \quad (13)$$

for some complex-valued function $h'(j_i - j'_i : i \in A)$. Comparing Eqs. (10) and (13), one concludes that $\langle \mathbf{k}'_{\text{encode}} | \mathcal{F}^\dagger \mathcal{F} | \mathbf{k}_{\text{encode}} \rangle = \delta_{\mathbf{k}, \mathbf{k}'} \Lambda_{\mathcal{F}, \mathcal{F}'}$ for some $\Lambda_{\mathcal{F}, \mathcal{F}'}$ independent of both \mathbf{k} and \mathbf{k}' . Thus, the QECC given in Eq. (8) corrects the phase shift errors as promised.

Conversely, from Eq. (13), one concludes that Eq. (8) corrects phase errors if and only if

$$\sum_{\{j, j'\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \left(\prod_{i \in U} \delta_{j_i, j'_i} \right) h'(j_i - j'_i : i \in A) \right] = \delta_{\mathbf{k}, \mathbf{k}'} \Lambda_{h'} \quad (14)$$

for *any* complex-valued function $h'(j_i - j'_i : i \in A)$. Hence, from Eq. (10), one concludes that Eq. (7) is able to correct additive spin flips errors. \square

In essence, Lemma 1 tells us that the abilities to correct additive spin flip and phase shift form a dual pair under the discrete Fourier transform of quantum registers. An interesting case occurs when $N = 2$. Here, additive spin flip is the only possible kind of spin flip error. As a result, the abilities to correct spin flip and phase shift errors in $N = 2$ form a dual pair under Lemma 1. And this special form of Lemma 1 was proven earlier by various authors (see, for example, Refs. [6, 7, 16]).

Corollary 1. *If a QECC handles both spin flip and phase shift errors on the same set of quantum registers, then this QECC handles any general quantum errors occurring at the same set of quantum registers.*

Proof. Combining Eqs. (10) and (12), one knows that Eq. (10) holds for *any* complex-valued function $g(j_i, j'_i : i \in A)$. By putting $\langle j'_i | \mathcal{E}_i^{\dagger} \mathcal{E}_i | j_i \rangle = g(j_i, j'_i)$ for all $i \in A$, then one concludes that the above QECC is capable of correcting any general quantum errors as promised. \square

Lemma 2. *Suppose QECCs $C1$ and $C2$ handle phase shift and spin flip errors, respectively, for the same set of quantum registers. Then, pasting the two codes together by first encodes the quantum state using $C1$ then further encodes the resultant quantum state using $C2$, one obtains a QECC C which corrects general errors in the same set of quantum registers.*

Proof. Clearly C can handle spin flip errors occurring at the specified quantum registers. So from Corollary 1 it remains to show that C corrects phase errors as well. Let the encodings for $C1$ and $C2$ be $|\mathbf{k}\rangle \mapsto \sum_{\{j\}} a_j^{(\mathbf{k})} |\mathbf{j}\rangle$ and $|\mathbf{j}\rangle \mapsto \sum_{\{\mathbf{p}\}} b_{\mathbf{p}}^{(\mathbf{j})} |\mathbf{p}\rangle$, respectively. Then using the same set of notations as in the proof of Lemma 1 one knows that

$$\begin{aligned} & \langle \mathbf{k}'_{\text{encode}} | \mathcal{F}'^\dagger \mathcal{F} | \mathbf{k}_{\text{encode}} \rangle \\ &= \sum_{\{j, j', \mathbf{p}, \mathbf{p}'\}} \left[\bar{a}_{j'}^{(\mathbf{k}')} a_j^{(\mathbf{k})} \bar{b}_{\mathbf{p}'}^{(\mathbf{j}')} b_{\mathbf{p}}^{(\mathbf{j})} \left(\prod_{i \in U} \delta_{p_i, p'_i} \right) \left(\prod_{i \in A} \langle p'_i | \mathcal{F}_i'^\dagger \mathcal{F}_i | p_i \rangle \right) \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{\{j, j', p, p'\}} \left\{ \bar{a}_{j'}^{(k')} a_j^{(k)} \bar{b}_{p'}^{(j')} b_p^{(j)} \delta_{p, p'} \left[\prod_{i \in A} g_i(p_i : i \in A) \right] \right\} \\
&= \sum_{\{j, j', p\}} \left\{ \bar{a}_{j'}^{(k')} a_j^{(k)} \bar{b}_p^{(j')} b_p^{(j)} \left[\prod_{i \in A} g(p_i : i \in A) \right] \right\} \quad (15)
\end{aligned}$$

for some complex-valued functions $g_i(p_i : i \in A)$ for all $i \in A$.

Since $C2$ handles spin flips, one demands that whenever $j \neq j'$,

$$\sum_{\{p\}} \bar{b}_p^{(j')} b_p^{(j)} = 0 = \sum_{\{p, p'\}} \bar{b}_{p'}^{(j')} b_p^{(j)} \langle p' | \mathcal{E} | p \rangle, \quad (16)$$

where \mathcal{E} denotes a possible spin flip error that can be handled by the QECC $C2$. Consequently,

$$\sum'_{\{p\}} \bar{b}_p^{(j')} b_p^{(j)} = 0, \quad (17)$$

where the above primed sum is over either (1) all the p that is affected by the error \mathcal{E} , or (2) all the p that is unaffected by the error \mathcal{E} .

From Eq. (17), it is easy to see that after summing over all p_i s in Eq. (15), one will arrive at

$$\langle k'_{\text{encode}} | \mathcal{F}^\dagger \mathcal{F} | k_{\text{encode}} \rangle = \sum_{\{j, j'\}} \left[\bar{a}_{j'}^{(k')} a_j^{(k)} \delta_{j, j'} \left(\prod_{i \in A} h_i(j_i : i \in A) \right) \right], \quad (18)$$

for some complex-valued function $h_i(j_i : i \in A)$. As $C1$ handles phase shift, one concludes that $\langle k'_{\text{encode}} | \mathcal{F}^\dagger \mathcal{F} | k_{\text{encode}} \rangle = \delta_{k, k'} \Lambda_{\mathcal{F}, \mathcal{F}'}$. Hence, the Lemma is proved. \square

At this point, I would like to remark that the proof of the abilities to correct both spin flip and phase shift implies the ability to correct a general error for $N = 2$ can be found in Refs. [6, 7, 14, 26]. Moreover, one should notice that the ordering of encoding in Lemma 2 is important. Encoding first using a spin flip code followed by a phase shift code does *not*, in general, result in a general QECC. After proving the above two technical lemmas, I report a method to construct QECCs from classical codes.

Theorem 2. *Suppose C is a classical (block or convolutional) code of rate r that corrects p (classical) errors for every q consecutive registers. Then, C can be extended to a QECC of rate r^2 that corrects at least p quantum errors for every q^2 consecutive quantum registers.*

Proof. Suppose C is a classical (block or convolutional) code. By mapping m to $|m\rangle$ for all $m \in \mathbb{Z}_N$, C can be converted to a quantum code for spin flip errors. Let C' be the QECC obtained by Fourier transforming each quantum register of C . Then Lemma 1 implies that C' is a code for phase shift errors. From Lemma 2, pasting codes C and C' together will create a QECC of rate r^2 . Finally, the fact that C' corrects at least p quantum errors for every q^2 consecutive quantum registers follows directly from Corollary 1. \square

Theorem 2 provides a powerful way to create high rate QECCs from high rate classical codes.

Example 2 ((Shor)). Starting with the simplest classical majority block code of rate $1/3$, namely, $|k\rangle \mapsto |k, k, k\rangle$ for $k = 0, 1$, Theorem 2 returns the famous Shor's single error correcting nine bit code [23] of rate $1/9$:

$$|k\rangle \mapsto \sum_{p,q,r=0}^1 (-1)^{k(p+q+r)} |p, p, p, q, q, q, r, r, r\rangle. \quad (19)$$

Alternatively, one may start with a high rate classical convolutional code. One of the simplest codes of this kind is the $1/2$ -rate code in Eq. (2). Being a non-systematic¹ and non-catastrophic² code (see, for example, chap. 4 in Ref. [13] for details), it serves as an ideal starting point to construct good QCCs. First, let me write down this code in quantum mechanical form:

Lemma 3. *The rate $1/2$ QCC*

$$\bigotimes_{i=1}^{+\infty} |k_i\rangle \mapsto |\mathbf{k}_{\text{encode}}\rangle \equiv \bigotimes_{i=1}^{+\infty} |k_i + k_{i-2}, k_i + k_{i-1} + k_{i-2}\rangle, \quad (20)$$

where $k_i \in \mathbb{Z}_N$ for all $i \in \mathbb{Z}^+$, can correct up to one spin flip error for every four consecutive quantum registers.

Proof. Here, I give a “quantum version” of the proof. Using notations in the proof of Theorem 1, I consider $\langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle$ again. Clearly, the worst case happens when errors \mathcal{E} and \mathcal{E}' occur at different quantum registers. And in this case, Eq. (20) implies that exactly two of the following four equations hold:

$$\begin{cases} k_{2i} + k_{2i-2} = k'_{2i} + k'_{2i-2} \\ k_{2i} + k_{2i-1} + k_{2i-2} = k'_{2i} + k'_{2i-1} + k'_{2i-2} \\ k_{2i+1} + k_{2i-1} = k'_{2i+1} + k'_{2i-1} \\ k_{2i+1} + k_{2i} + k_{2i-1} = k'_{2i+1} + k'_{2i} + k'_{2i-1} \end{cases} \quad (21)$$

for all $i \in \mathbb{Z}^+$. One may regard k_i s as unknowns and k'_i s as arbitrary but fixed constants. Then, by straight forward computation, one can show that picking *any* two equations out of Eq. (21) for each i will form an invertible system with the unique solution $k_i = k'_i$ for all $i \in \mathbb{Z}^+$. Therefore, $\langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle = \delta_{\mathbf{k}, \mathbf{k}'} \delta_{\mathcal{E}, \mathcal{E}'}$ and hence this lemma is proved. \square

Example 3. Theorem 2 and Lemma 3 imply that

$$\bigotimes_{i=1}^{+\infty} |k_i\rangle \mapsto \bigotimes_{i=1}^{+\infty} \left[\sum_{p_1, q_1, \dots} \frac{1}{N} \omega_N^{(k_i + k_{i-2})p_i + (k_i + k_{i-1} + k_{i-2})q_i} |p_i + p_{i-1}, p_i + p_{i-1} + q_{i-1}, q_i + q_{i-1}, q_i + q_{i-1} + p_i\rangle \right], \quad (22)$$

¹ That is, both b_i and c_i are not equal to a_i .

² That is, a finite number of channel errors does not create an infinite number of decoding errors.

where $k_i \in \mathbb{Z}_N$ for all $i \in \mathbb{Z}^+$, is a rate $1/4$ QCC capable of correcting up to one quantum error for every sixteen consecutive quantum registers.

In what follows, I show that this code can in fact correct up to one quantum error per every eight consecutive quantum registers.

Proof. Suppose \mathcal{E} and \mathcal{E}' be two quantum errors affecting at most one quantum register per every eight consecutive ones. By considering $\langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle$, I know that at least six of the following eight equations hold:

$$\left\{ \begin{array}{l} p_{2i-1} + p_{2i-2} = p'_{2i-1} + p'_{2i-2} \\ p_{2i-1} + p_{2i-2} + q_{2i-2} = p'_{2i-1} + p'_{2i-2} + q'_{2i-2} \\ q_{2i-1} + q_{2i-2} = q'_{2i-1} + q'_{2i-2} \\ q_{2i-1} + q_{2i-2} + p_{2i-1} = q'_{2i-1} + q'_{2i-2} + p'_{2i-1} \\ p_{2i} + p_{2i-1} = p'_{2i} + p'_{2i-1} \\ p_{2i} + p_{2i-1} + q_{2i-1} = p'_{2i} + p'_{2i-1} + q'_{2i-1} \\ q_{2i} + q_{2i-1} = q'_{2i} + q'_{2i-1} \\ q_{2i} + q_{2i-1} + p_{2i} = q'_{2i} + q'_{2i-1} + p'_{2i} \end{array} \right. \quad (23)$$

for all $i \in \mathbb{Z}^+$. Again, I regard p_i and q_i as unknowns; and p'_i and q'_i as arbitrary but fixed constants. Then, it is straight forward to show that choosing *any* six equations in Eq. (23) for each $i \in \mathbb{Z}^+$ would result in a consistent system having a unique solution of $p_i = p'_i$ and $q_i = q'_i$ for all $i \in \mathbb{Z}^+$. Consequently,

$$\begin{aligned} & \langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle \\ &= \sum_{\{\mathbf{p}, \mathbf{q}\}} \left\{ \prod_{i=1}^{+\infty} \left[\omega_N^{\sum_{j=2i-1}^{2i} p_j (k_j + k_{j-2} - k'_j - k'_{j-2}) + q_j (k_j + k_{j-1} + k_{j-2} - k'_j - k'_{j-1} - k'_{j-2})} \right. \right. \\ & \quad \left. \left. \times \langle f_i | \mathcal{E}'^\dagger | f_i \rangle \langle g_i | \mathcal{E} | g_i \rangle \right] \right\} \end{aligned} \quad (24)$$

for some linearly independent functions $f_i(\mathbf{p}, \mathbf{q})$ and $g_i(\mathbf{p}, \mathbf{q})$.

Now, I consider a basis $\{h_i(\mathbf{p}, \mathbf{q})\}$ for the orthogonal complement of the span of $\{f_i, g_i\}_{i \in \mathbb{Z}^+}$. By summing over all h_i s while keeping f_i s and g_i s constant in Eq. (24), one ends up with the constraints that $k_i = k'_i$ for all $i \in \mathbb{Z}^+$. Thus,

$$\begin{aligned} & \langle \mathbf{k}'_{\text{encode}} | \mathcal{E}'^\dagger \mathcal{E} | \mathbf{k}_{\text{encode}} \rangle \\ &= \delta_{\mathbf{k}, \mathbf{k}'} \sum_{\{\mathbf{p}, \mathbf{q}\}} \left[\prod_{i=1}^{+\infty} (\langle f_i(\mathbf{p}, \mathbf{q}) | \mathcal{E}'^\dagger | f_i(\mathbf{p}, \mathbf{q}) \rangle \langle g_i(\mathbf{p}, \mathbf{q}) | \mathcal{E} | g_i(\mathbf{p}, \mathbf{q}) \rangle) \right]. \end{aligned} \quad (25)$$

Hence, Eq. (22) corrects up to one quantum error per every eight consecutive quantum registers. \square

From the discussion following Example 1, the encoding in Eq. (20) can be done efficiently with the help of reversible computation [10, 21, 22].

4 Outlook

It is instructive to investigate the coding ability of QCCs as compared to that of QBCs. Knill and Laflamme [17] proved that it is impossible to construct a four qubit QBC that corrects one general quantum error. Their result can be extended to the case when $N > 2$ [10]. Here, with a slight modification of Knill and Laflamme's proof, I show that:

Theorem 3. *It is not possible to construct a QCC which corrects one general quantum error for every four consecutive quantum registers.*

Proof. Clearly, the QCC must be of rate $1/4$. And with a simple permutation of the quantum registers, a general QCC of rate $1/4$ can be written as

$$|k\rangle \longmapsto |k_{\text{encode}}\rangle \equiv \sum_{\{w,x,y,z\}} a_{w,x,y,z}^{(k)} |w,x,y,z\rangle. \quad (26)$$

Without loss of generality, I may assume that quantum errors occur in *any* one of the following four set of registers: $|w\rangle$, $|x\rangle$, $|y\rangle$ and $|z\rangle$.

Then, following Knill and Laflamme [17] by considering the action of errors in the above four sets of registers, one arrives at $\rho^{(k)} \rho^{(k')} = 0$ and $\rho^{(k)} = \rho^{(k')}$ for all $k \neq k'$. where

$$\rho_{w',x';w,x}^{(k)} = \sum_{\{y,z\}} \bar{a}_{w',x',y,z}^{(k)} a_{w,x,y,z}^{(k)}. \quad (27)$$

Hence, the reduced (Hermitian) density matrices $\rho^{(k)}$ are nilpotent for all k . This is possible only if $a_{w,x,y,z}^{(k)} = 0$ for all k, w, x, y, z . This contradicts the assumption that $|k_{\text{encode}}\rangle$ is a QCC. \square

It is, however, unclear if QCC can perform better than QBC in other situations. And further investigation along this line is required.

Acknowledgments: I would like to thank T. M. Ko for introducing me the subject of convolutional codes. I would also like to thank Debbie Leung, H.-K. Lo and Eric Rains for their useful discussions. This work is supported by the Hong Kong Government RGC grant HKU 7095/97P.

References

1. Bennett, C. H.: Logical reversibility of computation. IBM J. Res. Dev. **17** (1973) 525–532
2. Bennett, C. H.: Time / Space trade-offs for reversible computation. SIAM J. Comp. **18** (1989) 766–776
3. Bennett, C. H., DiVincenzo, D. P., Smolin, J. A., Wootters, W. K.: Mixed-state entanglement and quantum error correction. Phys. Rev. A **54** (1996) 3824–3851
4. Braunstein, S. L.: Error correction for continuous quantum variables. Los Alamos electronic preprint archive [quant-ph/9711049](https://arxiv.org/abs/quant-ph/9711049)

5. Braunstein, S. L., Smolin, J. A.: Perfect quantum-error-correcting coding in 24 laser pulses. *Phys. Rev. A* **55** (1997) 945–950
6. Calderbank, A. R., Rains, E. M., Shor, P. W., Sloane, N. J. A.: Quantum error correction via codes over $GF(4)$. *IEEE Trans. Inf. Theo.*, to appear. Also available from Los Alamos electronic preprint archive **quant-ph/9608006**
7. Calderbank, A. R., Rains, E. M., Shor, P. W., Sloane, N. J. A.: Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.* **78** (1997) 405–408
8. Calderbank, A. R., Shor, P. W.: Good quantum error-correcting codes exist. *Phys. Rev. A* **54** (1996) 1098–1105
9. Chau, H. F.: Correcting quantum errors in higher spin systems. *Phys. Rev. A* **55** (1997) 839–841
10. Chau, H. F.: Five quantum register error correction code for higher spin systems. *Phys. Rev. A* **56** (1997) 1–4
11. Chau, H. F.: Quantum convolutional codes. Los Alamos electronic preprint archive **quant-ph/9712029** (1997)
12. Chau, H. F., Lo, H.-K.: One-way functions in reversible computations. *Cryptologia* **21** (1997) 139–148
13. Dholakia, A.: Introduction to convolutional codes with applications. Kluwer, Dordrecht (1994)
14. Ekert, A., Macchiavello, C.: Quantum error correction for communication. *Phys. Rev. Lett.* **77** (1996) 2585–2588
15. Gottesman, D.: Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev. A* **54** (1996) 1862–1868
16. Knill, E.: Non-binary unitary error bases and quantum codes. Los Alamos electronic preprint archive **quant-ph/9608048**
17. Knill, E., Laflamme, R.: Theory of quantum error-correcting codes. *Phys. Rev. A* **55** (1997) 900–911
18. Laflamme, R., Miquel, C., Paz, J. P., Zurek, W. H.: Perfect quantum error correcting code. *Phys. Rev. Lett.* **77** (1996) 198–201
19. Lloyd, S., Slotine, J.-J. E.: Analog quantum error correction. Los Alamos electronic preprint archive **quant-ph/9711021**
20. Piret, Ph.: Convolutional codes: an algebraic approach. MIT Press, Cambridge, MA (1988)
21. Rains, E. M.: Nonbinary quantum codes. Los Alamos electronic preprint archive **quant-ph/9703048**
22. Rains, E. M., Hardin, R. H., Shor, P. W., Sloane, N. J. A.: Nonadditive quantum code. *Phys. Rev. Lett.* **79** (1997) 953–954
23. Shor, P. W.: Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52** (1995) 2493–2496
24. Steane, A. M.: Simple quantum error-correcting codes. *Phys. Rev. A* **54** (1996) 4741–4751
25. Steane, A. M.: Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77** (1996) 793–797
26. Steane, A.: Multiple-particle interference and quantum error correction. *Proc. Roy. Soc. Lond. A* **452** (1996) 2551–2577

On the Existence of Nonadditive Quantum Codes

Vwani P. Roychowdhury and Farrokh Vatan

Electrical Engineering Department
UCLA

Los Angeles, CA 90095

vwani@ee.ucla.edu and vatan@ee.ucla.edu

Abstract. Most of the quantum error-correcting codes studied so far fall under the category of additive (or stabilizer) quantum codes, which are closely related to classical linear codes. The existence and general constructions of efficient quantum codes that do not have such an underlying structure have remained elusive. Recently, specific examples of nonadditive quantum codes with minimum distance 2 have been presented. We, however, show that there exist infinitely many non-trivial nonadditive codes with different minimum distances, and high rates. In fact, we show that nonadditive codes that correct t errors can reach the asymptotic rate $R = 1 - 2H_2(2t/n)$, where $H_2(x)$ is the binary entropy function. In the process, we also develop a general set of sufficient conditions for a quantum code to be nonadditive. Finally, we introduce the notion of *strongly* nonadditive codes, and provide a construction for an $((11,2,3))$ strongly nonadditive code.

key words: quantum code, additive code, Gilbert–Varshamov bound.

1 Introduction

Almost all quantum error-correcting codes known so far are additive (or stabilizer) codes. An additive code can be described as follows. Consider the group \mathcal{G} of unitary operators on the Hilbert space \mathbb{C}^{2^n} defined by the tensor products $\pm M_1 \otimes M_2 \otimes \cdots \otimes M_n$, where each M_i is either the identity $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or one of the Pauli matrices σ_x , σ_z , or $\sigma_y = i\sigma_x\sigma_z$. Then an additive code is a subspace \mathcal{Q} of \mathbb{C}^{2^n} for which there is an Abelian subgroup H of \mathcal{G} such that every vector of \mathcal{Q} is a fixed point of every operator in H [3,47].¹ This approach leads to a close connection between self-orthogonal (under a specific inner product) linear binary codes and additive codes, such that the minimum distance of the additive code is determined from the binary code.

¹ This is actually the definition of a *real* additive code; i.e., a code which has a basis consisting of vectors from \mathbb{R}^{2^n} . In this paper we restrict ourselves to the set of real codes, but this does not restrict our results, since every additive code is equivalent to a real one [11].

It is natural to ask whether there is any quantum error-correcting code that cannot be constructed in this way, directly or via some equivalence. We should make here a comment on the correct formulation of this question. Since the dimension of every additive quantum code is a power of 2, any quantum code whose dimension is not a power of 2 is not additive or equivalent to an additive code; especially, any subspace of an additive code with dimension not a power of 2 is a nonadditive code. We call such codes *trivial nonadditive codes*. But we prove a general theorem that shows that infinite families of non-trivial nonadditive codes with different values of d exist. The nonadditiveness of these codes does not follow from their dimensions (the dimensions of these codes are also powers of two), but from their very special structure. Moreover, we show that these nonadditive codes asymptotically reach the same rate as Calderbank–Shor–Steane codes.

We also propose the notion of *strongly nonadditive* codes: a quantum code \mathcal{Q} is strongly nonadditive if the trivial code \mathbb{C}^{2^n} is the only additive code that contains any code equivalent to \mathcal{Q} . Now the interesting problem is to find strongly nonadditive quantum codes. Recently in [13] it is shown that a $((5, 6, 2))$ strongly nonadditive code exists, which is better than any $((5, K, 2))$ additive code. Later in [12], Rains showed that there exists a $((2m, 4^{m-1}, 2))$ nonadditive code, for every $m \geq 3$. We present an $((11, 2, 3))$ strongly nonadditive code.

In Section 3 we first determine a criterion that guarantees additiveness and strongly nonadditiveness of quantum codes, and then we present our examples of additive and strongly nonadditive codes.

2 Preliminaries

Consider the Hilbert space \mathbb{C}^{2^n} with its standard basis $|v_1\rangle, \dots, |v_{2^n}\rangle$, where v_1, \dots, v_{2^n} is a list of binary vectors of length n in $\{0, 1\}^n$. For every binary vector α of length n , we define the unitary operators X_α and Z_α by the following equations

$$\begin{aligned} X_\alpha |v_i\rangle &= |v_i + \alpha\rangle, \\ Z_\alpha |v_i\rangle &= (-1)^{v_i \cdot \alpha} |v_i\rangle. \end{aligned}$$

Note that $X_\alpha Z_\beta = (-1)^{\alpha \cdot \beta} Z_\beta X_\alpha$.

Let \mathcal{G} be the group of all unitary operators of the form $\pm M_1 \otimes \dots \otimes M_n$, where $M_i \in \{I, \sigma_x, \sigma_y, \sigma_z\}$. Then every member of \mathcal{G} can be represented uniquely as $(-1)^\lambda X_\alpha Z_\beta$, where $\lambda \in \{0, 1\}$ and $\alpha, \beta \in \{0, 1\}^n$. For every subgroup \mathcal{S} of \mathcal{G} , let $\overline{\mathcal{S}} \subset \{0, 1\}^{2n}$ be the set of all vectors $(\alpha|\beta)$ such that either $X_\alpha Z_\beta \in \mathcal{S}$ or $-X_\alpha Z_\beta \in \mathcal{S}$. We say $\overline{\mathcal{S}}$ is *totally singular* if for every $(\alpha|\beta) \in \overline{\mathcal{S}}$ we have $\alpha \cdot \beta = 0$. We also define a special inner product on $\{0, 1\}^{2n}$ as

$$((a|b), (a'|b')) = a \cdot b' + a' \cdot b, \quad (1)$$

where the right-hand side is evaluated in $\text{GF}(2)$. For any quantum code \mathcal{Q} in \mathbb{C}^{2^n} , we define the *stabilizer* $\mathcal{H}_{\mathcal{Q}}$ of \mathcal{Q} as

$$\mathcal{H}_{\mathcal{Q}} = \{ \varphi \in \mathcal{G} : \varphi |x\rangle = |x\rangle \text{ for every } |x\rangle \text{ in } \mathcal{Q} \}.$$

Then it is easy to check that \mathcal{H}_Q is an Abelian group and every element of \mathcal{H}_Q squares to the identity operator. So $\overline{\mathcal{H}_Q}$ is totally singular. It also follows that \mathcal{H}_Q is isomorphic to a vector space $\text{GF}(2)^m$, for some m . This means that \mathcal{H}_Q is generated by operators $\varphi_1, \dots, \varphi_m \in \mathcal{H}_Q$ and every $\varphi \in \mathcal{H}_Q$ can be written (uniquely, up to the order of the φ_i 's) as $\varphi = \varphi_1^{c_1} \dots \varphi_m^{c_m}$, where $c_i \in \{0, 1\}$. In this case the quantum code Q has dimension 2^{n-m} . Suppose that $\varphi_i = (-1)^{\lambda_i} X_{\alpha_i} Z_{\beta_i}$. So $\overline{\mathcal{H}_Q}$ can be determined by its $m \times (2n)$ binary generating matrix

$$M = \left(\begin{array}{c|c} \alpha_1 & \beta_1 \\ \vdots & \vdots \\ \alpha_m & \beta_m \end{array} \right). \quad (2)$$

Note that if such a matrix M obtained from a stabilizer, then $\alpha_i \cdot \beta_i = 0$ and $\alpha_i \cdot \beta_j + \alpha_j \cdot \beta_i = 0$, for every i and j . A quantum code Q is called *additive* (or *stabilizer*) if it is defined by its stabilizer \mathcal{H}_Q , i.e.,

$$Q = \left\{ |x\rangle \in \mathbb{C}^{2^n} : \varphi |x\rangle = |x\rangle \text{ for every } \varphi \in \mathcal{H}_Q \right\}.$$

The quantum codes Q_1 and Q_2 in \mathbb{C}^{2^n} are *locally equivalent* if there is a transversal operator $\mathcal{U} = u_1 \otimes \dots \otimes u_n$, with $u_i \in \text{SU}(2)$, mapping Q_1 into Q_2 . We say these codes are *globally equivalent*, or simply *equivalent*, if Q_1 is locally equivalent to a code obtained from Q_2 by a permutation on qubits.

A quantum code $Q \subseteq \mathbb{C}^{2^n}$ is called **nonadditive** if it is not equivalent to any additive code; moreover, Q is **strongly nonadditive** if the only additive code that contains any code equivalent to Q is the trivial code \mathbb{C}^{2^n} ; in other words, if $\pm X_\alpha Z_\beta$ is in the stabilizer of any code equivalent to a supercode of Q then $\alpha = \beta = \mathbf{0}$.

A K -dimensional subspace of \mathbb{C}^{2^n} that as an error-correcting quantum code can protect against $< d/2$ errors, is called an $((n, K, d))$ code. If this code is additive, then $K = 2^k$, for some k , and is called an $[[n, k, d]]$ code. The following theorem gives a sufficient condition for a subspace of \mathbb{C}^{2^n} to be an $((n, K, d))$ code. Here $\text{wt}(c)$ denotes the Hamming weight of the binary vector c , i.e., the number of 1-components in c , and $\alpha \cup \beta$ is the binary vector resulting from a component-wise OR operation of α and β ; for example $(10110) \cup (00101) = (10111)$.

Theorem 1. ([1], [8]) *Let Q be a K -dimensional subspace of \mathbb{C}^{2^n} . Consider an orthonormal basis for Q of the form $\{|c_i\rangle : i = 1, \dots, K\}$. Then Q is an $((n, K, d))$ code if $\langle c_i | X_\alpha Z_\beta | c_j \rangle = 0$ for every $1 \leq i, j \leq K$ and for every $\alpha, \beta \in \{0, 1\}^n$ with $1 \leq \text{wt}(\alpha \cup \beta) \leq d-1$. In general, a necessary and sufficient condition for Q to be an $((n, K, d))$ code is that for all $1 \leq i, j \leq K$ and $\text{wt}(\alpha \cup \beta) \leq d-1$ we have $\langle c_i | X_\alpha Z_\beta | c_i \rangle = \langle c_j | X_\alpha Z_\beta | c_j \rangle$ and if $i \neq j$ then $\langle c_i | X_\alpha Z_\beta | c_j \rangle = 0$.*

□

For an additive code Q with stabilizer \mathcal{H}_Q there is a sufficient condition in term of the dual of \mathcal{H}_Q with respect to the inner product defined by equation (II) for Q to be a t -error-correcting code.

Theorem 2. ([3], [7]) Let \mathcal{Q} be an additive code with stabilizer $\mathcal{H}_{\mathcal{Q}}$. Let $\overline{\mathcal{H}_{\mathcal{Q}}}^{\perp}$ be the space orthogonal to $\mathcal{H}_{\mathcal{Q}}$ with respect to the inner product (11). If for every pair of binary vectors $\alpha, \beta \in \{0, 1\}^n$ with $\text{wt}(\alpha \cup \beta) \leq d-1$ we have $(\alpha|\beta) \notin \overline{\mathcal{H}_{\mathcal{Q}}}^{\perp} \setminus \overline{\mathcal{H}_{\mathcal{Q}}}$ then \mathcal{Q} is an $[[n, k, d]]$ additive code. \square

3 Existence of Nonadditive Codes

3.1 Quantum Codes Equivalent to Additive Codes

We study the quantum codes equivalent to additive codes. For such code \mathcal{Q} , we find a sufficient condition that guarantees that the stabilizer of \mathcal{Q} contains a nontrivial operator.

We begin with some useful notions and notations. Let $|c_1\rangle, \dots, |c_{2^n}\rangle$ be the standard orthonormal basis of \mathbb{C}^{2^n} , where each c_i is a binary vector of length n .

For the vector $|x\rangle = \sum_{i=1}^{2^n} \lambda_i |c_i\rangle$, we define the *support* of $|x\rangle$ as

$$\text{supp}(|x\rangle) = \{c_i \in \{0, 1\}^n : \lambda_i \neq 0\}.$$

Let $\mathcal{C} \subseteq \{0, 1\}^n$ be a set of binary vectors. Define the vector $|\mathcal{C}\rangle$ in \mathbb{C}^{2^n} as

$$|\mathcal{C}\rangle = \frac{1}{|\mathcal{C}|^{1/2}} \sum_{c \in \mathcal{C}} |c\rangle.$$

(If \mathcal{C} is empty then $|\mathcal{C}\rangle$ is the zero vector.) For any binary vector α of length $m < n$, define

$$\mathcal{C}_{\alpha} = \{x \in \{0, 1\}^{n-m} : (\alpha, x) \in \mathcal{C}\}. \quad (3)$$

So to construct \mathcal{C}_{α} , consider all vectors in \mathcal{C} starting with α (if there is any), then delete α from these vectors. Note that \mathcal{C}_{α} may be empty.

For a quantum code \mathcal{Q} , let us define the *generalized stabilizer* of \mathcal{Q} as the set $GS(\mathcal{Q})$ of all unitary operators \mathcal{V} on \mathbb{C}^{2^n} such that $\mathcal{V}|x\rangle = |x\rangle$ for every $|x\rangle \in \mathcal{Q}$. Then the *stabilizer* of \mathcal{Q} is $\text{St}(\mathcal{Q}) = \mathcal{G} \cap GS(\mathcal{Q})$.

Lemma 1. Suppose that the quantum codes \mathcal{Q}_1 and \mathcal{Q}_2 are locally equivalent via the transversal unitary operator \mathcal{U} . Then for every $M \in GS(\mathcal{Q}_1)$ the operator $\mathcal{U}M\mathcal{U}^{\dagger}$ is in $GS(\mathcal{Q}_2)$.

Proof. Let $|x\rangle \in \mathcal{Q}_2$. Now, we know that there exists a code word $|y\rangle \in \mathcal{Q}_1$ such that $|x\rangle = \mathcal{U}|y\rangle$. Since $M|y\rangle = |y\rangle$, so $(M\mathcal{U}^{\dagger})\mathcal{U}|y\rangle = |y\rangle$, and therefore $(\mathcal{U}M\mathcal{U}^{\dagger})\mathcal{U}|y\rangle = \mathcal{U}|y\rangle$. This implies $(\mathcal{U}M\mathcal{U}^{\dagger})|x\rangle = |x\rangle$. \square

We are interested in the case of $M \in \mathcal{G}$, i.e., $M = M_1 \otimes \dots \otimes M_n$, where $M_j \in \{I, \sigma_x, \sigma_y, \sigma_z\}$. We define $\text{wt}(M)$ the *weight* of any $M \in \mathcal{G}$ as the number of

j 's such that $M_j \neq I$. In this case $\mathcal{U}\mathcal{M}\mathcal{U}^\dagger = v_1 \otimes \cdots \otimes v_n$ such that $\det(v_j) = \pm 1$ and if $M_j = I$ then $v_j = I$, otherwise

$$v_j = \eta_j \begin{pmatrix} a_j & b_j \\ \pm b_j^* & -a_j \end{pmatrix}, \quad \eta_j \in \{1, i\}, \quad a_j \in \mathbb{R} \text{ and } b_j \in \mathbb{C}. \quad (4)$$

If $\mathcal{U} \in \text{SU}(2)^{\otimes n}$ then \mathcal{U} is of the form $u_1 \otimes \cdots \otimes u_n$, where each u_j is defined by a matrix of the form

$$\begin{pmatrix} e^{i\alpha} \cos \theta & e^{i\beta} \sin \theta \\ -e^{-i\beta} \sin \theta & e^{-i\alpha} \cos \theta \end{pmatrix}. \quad (5)$$

If $M_j = \sigma_x$, σ_z or σ_y , then the corresponding v_j , respectively, is

$$\left. \begin{aligned} & \begin{pmatrix} \sin 2\theta \cos(\alpha - \beta) & \cos^2 \theta e^{i2\alpha} - \sin^2 \theta e^{i2\beta} \\ \cos^2 \theta e^{-i2\alpha} - \sin^2 \theta e^{-i2\beta} & -\sin 2\theta \cos(\alpha - \beta) \end{pmatrix}, \\ & \begin{pmatrix} \cos 2\theta & -\sin 2\theta e^{i(\alpha+\beta)} \\ -\sin 2\theta e^{-i(\alpha+\beta)} & -\cos 2\theta \end{pmatrix}, \\ \text{or } & \begin{pmatrix} -i \sin 2\theta \sin(\alpha - \beta) & -\cos^2 \theta e^{i2\alpha} - \sin^2 \theta e^{i2\beta} \\ \cos^2 \theta e^{-i2\alpha} + \sin^2 \theta e^{-i2\beta} & i \sin 2\theta \sin(\alpha - \beta) \end{pmatrix}. \end{aligned} \right\} \quad (6)$$

We call a matrix v_i as (4) *full* if $a_i \cdot b_i \neq 0$; and we say the unitary operator $\mathcal{V} = v_1 \otimes \cdots \otimes v_n$ is *thin* if none of v_i 's is full. In the next proof we will use this property that if \mathcal{V} is thin then $|\text{supp}(\mathcal{V}|x\rangle)| = |\text{supp}(|x\rangle)|$, for every $|x\rangle$.

A quantum code \mathcal{Q} is called *real* if \mathcal{Q} has a basis consisting of real vectors; i.e., if $|x\rangle = \sum_{i=1}^{2^n} \lambda_i |c_i\rangle$ is any vector in the basis, then $\lambda_i \in \mathbb{R}$, for every i .

An (n, K, d) binary code is a set $\mathcal{C} \subseteq \{0, 1\}^n$ of size K such that any two vectors in \mathcal{C} differ in at least d places, and d is the largest number with this property. Note that an $[n, k, d]$ binary linear code is an $(n, 2^k, d)$ binary code.

Theorem 3. *Suppose that the quantum codes \mathcal{Q}_1 and \mathcal{Q}_2 are locally equivalent via the transversal operator \mathcal{U} , \mathcal{Q}_2 is real and \mathcal{Q}_2 contains $|\mathcal{C}\rangle$, where \mathcal{C} is an (n, K, d) binary code with $d > k = \lceil \log_2 K \rceil$. Then the following claims hold.*

(i) *The image of $\text{St}(\mathcal{Q}_1)$ under the mapping $M \mapsto \mathcal{U}\mathcal{M}\mathcal{U}^\dagger$, which we call Γ , consists only of unitary operators of the form $\pm X_\alpha T$, where T is a Z -type unitary operator of the form*

$$T = \bigotimes_{j=1}^n \begin{pmatrix} e^{i\theta_j} & 0 \\ 0 & \pm e^{-i\theta_j} \end{pmatrix}. \quad (7)$$

(ii) *Let $\Delta = \{\alpha \in \{0, 1\}^n : \pm X_\alpha T \in \Gamma \text{ for some } T \text{ of the form (7)}\}$. Suppose that $\text{St}(\mathcal{Q}_2)$ does not contain any operator of the form $\pm X_0 Z_\beta$, with $\beta \neq \mathbf{0}$. Then $|\text{St}(\mathcal{Q}_1)| \leq |\Delta|$. \square*

The proof of this theorem can be found in [14].

We now present a criterion for nonadditiveness of quantum codes. First a useful notation. For a subset \mathcal{C} of $\{0, 1\}^n$ let

$$\mathcal{T}(\mathcal{C}) = \{x \in \{0, 1\}^n : x + \mathcal{C} \subseteq \mathcal{C}\}.$$

If \mathcal{C} is a binary *linear* code then $\mathcal{T}(\mathcal{C}) = \mathcal{C}$.

Theorem 4. *Suppose that the quantum code \mathcal{Q} of dimension 2^ℓ is real and contains $|\mathcal{C}\rangle$, where \mathcal{C} is an (n, K, d) binary code with $d > \lceil \log_2 K \rceil$. If the identity operator is the only unitary operator in the stabilizer of \mathcal{Q} and $2^{n-\ell} > |\mathcal{T}(\mathcal{C})|$ then \mathcal{Q} is nonadditive.*

Proof. Suppose, by contradiction, that \mathcal{Q} is equivalent to additive code \mathcal{Q}' via the transversal unitary operator \mathcal{U} which maps \mathcal{Q}' on \mathcal{Q} . Let Γ be the image of $\text{St}(\mathcal{Q}')$ under \mathcal{U} . Define $\Delta \subseteq \{0, 1\}^n$ as in (ii) of Theorem 3. Then $\Delta \subseteq \mathcal{T}(\mathcal{C})$. Thus

$$2^{n-\ell} = |\text{St}(\mathcal{Q}')| \leq |\Delta| \leq |\mathcal{T}(\mathcal{C})|,$$

which contradicts the assumption of the theorem. \square

When the binary code \mathcal{C} in the above theorem is linear we can formulate the theorem as follows.

Corollary 1. *Suppose that the quantum code \mathcal{Q} of dimension 2^ℓ is real and contains $|\mathcal{C}\rangle$, where \mathcal{C} is a linear $[n, k, d]$ code with $d > k$. If $\text{St}(\mathcal{Q}) = \{I\}$ and $n > k + \ell$ then \mathcal{Q} is nonadditive.* \square

Finally, we formulate a criterion that guarantees strongly nonadditiveness of quantum codes.

Theorem 5. *Suppose that the quantum code \mathcal{Q} is real and it contains $|\mathcal{C}\rangle$ where \mathcal{C} is an (n, K, d) binary code with $d > \lceil \log_2 K \rceil$. If $\text{St}(\mathcal{Q}) = \{I\}$ and $GS(\mathcal{Q})$ does not contain any operator of the form $X_\alpha T$, where $\alpha \neq \mathbf{0}$ and T is of the form (7), then \mathcal{Q} is strongly nonadditive.*

Proof. Suppose, by contradiction, that $\mathcal{Q} \subseteq \mathcal{Q}_1$ and $\mathcal{Q}_1 \neq \mathbb{C}^{2^n}$ is equivalent to an additive code \mathcal{Q}' with $\text{St}(\mathcal{Q}') \neq \{I\}$. Then, by Theorem 3, any nontrivial stabilizer φ of \mathcal{Q}' defines an operator $\mathcal{V} = v_1 \otimes \cdots \otimes v_n$ in $GS(\mathcal{Q}_1) \subseteq GS(\mathcal{Q})$, where $v_j = I$ or it is of the form (4) or (6). If all v_j have real matrices, then $\mathcal{V} \neq I$ and $\mathcal{V} \in \text{St}(\mathcal{Q})$, which is impossible. If at least one of v_j has a complex matrix, then \mathcal{V} is of the form $X_\alpha T$ with $\alpha \neq \mathbf{0}$, which is again impossible. \square

3.2 Construction of Nonadditive Codes

Examples of Nonadditive Codes. Now we show that there is an infinite family of nonadditive quantum error-correcting codes. These codes are constructed

following the scheme similar to the one described in Theorem 2.4 of [16]. Consider an $[n, k]$ binary code \mathcal{C} such that $\text{dist}(\mathcal{C})$ and $\text{dist}(\mathcal{C}^\perp)$ are both at least d_0 (\mathcal{C} needs not to be a weakly self-dual code).

First we define a function $\tau : \mathcal{C} \longrightarrow \{0, 1\}^n$ such that for $c, c' \in \mathcal{C}$ and $c \neq c'$ we have $\tau(c) + \tau(c') \notin \mathcal{C}^\perp$. This means $\tau(c)$ and $\tau(c')$ are in different cosets of \mathcal{C}^\perp in $\{0, 1\}^n$, for $c \neq c'$. Since there are 2^k different cosets, such mapping τ always can be defined.

Fix $d \leq d_0$, and let \mathcal{E} be the set of binary vectors of length n with weight $\leq d - 1$. Consider a subset $R = \{a_0, a_1, \dots, a_m\}$ of $\{0, 1\}^n$ such that $a_0 = \mathbf{0}$ and a_j is not of the form $c + a_i + e$, for $c \in \mathcal{C}$, $1 \leq i \leq j - 1$, and $e \in \mathcal{E}$. Then the vectors

$$|x_i\rangle = \sum_{c \in \mathcal{C}} (-1)^{\tau(c) \cdot a_i} |c + a_i\rangle \quad (8)$$

form a basis for a quantum code with distance d . To prove this, we show that $\langle x_i | X_\alpha Z_\beta | x_j \rangle = 0$, for $0 < \text{wt}(\alpha \cup \beta) < d$. The case $\alpha \neq \mathbf{0}$ or $i \neq j$ is straightforward. So we only consider the case $\alpha = \mathbf{0}$ and $i = j$. Then for $0 < \text{wt}(\beta) < d$ we have

$$\begin{aligned} \langle x_i | Z_\beta | x_i \rangle &= \left\langle \sum_{c \in \mathcal{C}} (-1)^{\tau(c) \cdot a_i} |c + a_i\rangle \left| \sum_{c \in \mathcal{C}} (-1)^{\tau(c) \cdot a_i + (c + a_i) \cdot \beta} |c + a_i\rangle \right. \right\rangle \\ &= (-1)^{a_i \cdot \beta} \sum_{c \in \mathcal{C}} (-1)^{c \cdot \beta} \\ &= 0. \end{aligned}$$

The last equality follows from the fact that $\text{dist}(\mathcal{C}^\perp) \geq d$, so $\beta \notin \mathcal{C}^\perp$.

Lemma 2. *In the above construction, suppose that*

$$(n - 1)2^k \sum_{i=0}^{d-1} \binom{n}{i} < 2^{n-1}. \quad (9)$$

Then it is possible to choose n linearly independent vectors a_1, a_2, \dots, a_n so that the $((n, n + 1, d))$ quantum code \mathcal{Q} with the basis $|x_0\rangle, |x_1\rangle, \dots, |x_n\rangle$ (each $|x_i\rangle$ is defined by (8)) has trivial stabilizer, i.e., $\text{St}(\mathcal{Q}) = \{I\}$.

Proof. Suppose that the vectors a_0, a_1, \dots, a_m with the desired properties are chosen. Then it is possible to choose a vector a_{m+1} such that a_1, \dots, a_m, a_{m+1} are independent and a_{m+1} is not of the form $c + a_i + e$ (for $c \in \mathcal{C}$, $1 \leq i \leq m$, and $e \in \mathcal{E}$) if $2^m + m \cdot 2^k \cdot \sum_{i=0}^{d-1} \binom{n}{i} < 2^n$. This shows that it is possible to choose n vector a_1, \dots, a_n with the desired properties.

Now we show that the identity operator is the only member of the stabilizer of \mathcal{Q} . Suppose that $X_\alpha Z_\beta$ is in the stabilizer of \mathcal{Q} . Since

$$X_\alpha Z_\beta |x_0\rangle = \sum_{c \in \mathcal{C}} (-1)^{c \cdot \beta} |c + \alpha\rangle$$

should be equal to $|x_0\rangle = \sum_{c \in \mathcal{C}} |c\rangle$ it follows that $\alpha \in \mathcal{C}$ and $\beta \in \mathcal{C}^\perp$. Similarly, for every $1 \leq i \leq n$ since

$$\begin{aligned} X_\alpha Z_\beta |x_i\rangle &= \sum_{c \in \mathcal{C}} (-1)^{\tau(c) \cdot a_i + (c+a_i) \cdot \beta} |c + a_i + \alpha\rangle \\ &= \sum_{c \in \mathcal{C}} (-1)^{\tau(c+\alpha) \cdot a_i + (c+a_i+\alpha) \cdot \beta} |c + a_i\rangle \\ &= \sum_{c \in \mathcal{C}} (-1)^{(\tau(c+\alpha)+\beta) \cdot a_i} |c + a_i\rangle \end{aligned}$$

should be equal to

$$|x_i\rangle = \sum_{c \in \mathcal{C}} (-1)^{\tau(c) \cdot a_i} |c + a_i\rangle,$$

it follows that $a_i \cdot (\tau(c) + \tau(c + \alpha) + \beta) = 0$, for every $1 \leq i \leq n$. Since a_i 's are independent, therefore $\tau(c) + \tau(c + \alpha) = \beta \in \mathcal{C}^\perp$, hence $\alpha = \mathbf{0}$. Now the conditions $a_i \cdot \beta = 0$ (for $1 \leq i \leq n$) imply $\beta = \mathbf{0}$. \square

Theorem 6. *Suppose that \mathcal{C} is an $[n, k, d_0]$ binary linear code such that $d_0 > k$ and $\text{dist}(\mathcal{C})$ and $\text{dist}(\mathcal{C}^\perp)$ are at least d . Moreover, suppose that n, k and d satisfy (9). Let ℓ be the greatest integer such that $2^\ell \leq 2^{n-k} / \sum_{i=0}^{d-1} \binom{n}{i}$. Suppose that $k + \ell < n$. Then there is a $((n, 2^\ell, d))$ nonadditive code.*

Proof. Consider the $((n, n+1, d))$ code \mathcal{Q}_0 constructed in the previous lemma. Then by Theorem 4.2 of [16] it is possible to add at least $2^\ell - (n+1)$ more vectors to \mathcal{Q}_0 to build an $((n, 2^\ell, d))$ code \mathcal{Q} , which is, by Corollary 1, nonadditive. \square

As an application we show that there are $((n, \lfloor 2^{n-1}/(n+1) \rfloor, 2))$ nonadditive codes, for every $n \geq 8$. Consider the $[n, 1, n]$ binary code $\mathcal{C} = \{\mathbf{0}, \mathbf{1}\}$. Then \mathcal{C}^\perp consists of all even weight vectors in $\{0, 1\}^n$, so it is an $[n, n-1, 2]$ code. The condition (9) satisfies if $n \geq 8$. Then by applying the above theorem (for $k = 1$ and $\ell = \lceil n - 1 - \log_2(n+1) \rceil$) we get the desired code. Other classes of binary codes for which the minimum distance of the code and its dual are known (such as Hamming codes and Reed–Muller codes) can be used to get nonadditive codes with different parameters.

Finally, we show that the nonadditive codes are almost as good as Calderbank–Shor–Steane (CSS) codes, at least in the case that the dimension of code is large enough.

To utilize the CSS codes for constructing nonadditive codes, we must modify them such that the new codes have trivial stabilizers. Let \mathcal{Q} be an $[[n, n-2k, d]]$ CCS code based on the weakly self-dual $[n, k]$ code \mathcal{C} with $\text{dist}(\mathcal{C}^\perp) \geq d$. Consider the basis for \mathcal{Q} consisting of vectors of the form $|x_a\rangle = \sum_{c \in \mathcal{C}} |c + a\rangle$, for $a \in \mathcal{C}^\perp/\mathcal{C}$.

Also consider the function $\tau: \mathcal{C} \rightarrow \{0, 1\}^n$ defined at the beginning of this section. We define the quantum code $\widehat{\mathcal{Q}}$ with basis

$$|y_a\rangle = \sum_{c \in \mathcal{C}} (-1)^{\tau(c) \cdot a} |c + a\rangle, \quad (10)$$

for $a \in \mathcal{C}^\perp / \mathcal{C}$. Then it is easy to check that $\widehat{\mathcal{Q}}$ is also an $[[n, n - 2k, d]]$ code.

Theorem 7. *Suppose that \mathcal{C} is an $[n, k, d_0]$ weakly self-dual binary code, and \mathcal{C}^\perp is an $[n, n - k, d_1]$ code. Assume $d_0 \geq k$ and $2^{n-2k-1} > n - k - 1$ (for example it is enough to have $k < (n - \log_2 n)/2$). For any $d \leq d_1$ that satisfies*

$$(2^{n-k} + (k-1)2^k) \sum_{i=0}^{d-1} \binom{n}{i} < 2^{n-1}, \quad (11)$$

we have an $((n, 2^{n-2k}, d))$ nonadditive code.

Proof. Let \mathcal{Q}_0 be the $[[n, n - 2k, d]]$ CSS code based on \mathcal{C} , and let $\widehat{\mathcal{Q}}_0$ be the quantum code obtained from \mathcal{Q}_0 as described in the preceding procedure. We can choose independent vectors a_1, \dots, a_{n-k} in \mathcal{C}^\perp such that a_i 's belong to different cosets of \mathcal{C} in \mathcal{C}^\perp . This is possible because $2^{n-2k-1} > n - k - 1$. We consider $|y_{a_1}\rangle, \dots, |y_{a_{n-k}}\rangle$ (defined by (10)) as vectors in $\widehat{\mathcal{Q}}_0$. Then we choose vectors a_{n-k+1}, \dots, a_n such that a_1, \dots, a_n are n independent vectors, and $\mathcal{Q}' = \widehat{\mathcal{Q}}_0 \cup \{|x_{a_{n-k+1}}\rangle, \dots, |x_{a_n}\rangle\}$, is an $((n, 2^{n-2k} + k, d))$ code. The inequality (11) implies that it is possible to choose a_{n-k+1}, \dots, a_n with the desired properties. Then the proof of Lemma 2 shows that $\text{St}(\mathcal{Q}') = \{I\}$.

Let \mathcal{Q} be the quantum code obtained from \mathcal{Q}' by removing any k vectors except $|y_{a_i}\rangle$, $i = 1, \dots, n$. Then $\text{St}(\mathcal{Q}) = \{I\}$ (because \mathcal{Q} contains the $|y_{a_i}\rangle$, $i = 1, \dots, n$). So, by Corollary 1 with $\ell = n - 2k$, \mathcal{Q} is nonadditive. \square

To show that there are weakly self-dual codes \mathcal{C} that satisfy the requirements of the above theorem, it is possible to apply the greedy method used in classical coding theory (see [10], Chap. 17). The same method is used in [5] to prove the existence of CSS codes meeting the Gilbert–Varshamov bound. This method gives the following bound.

Theorem 8. *For $d < \lambda n$, where $\lambda = H_2^{-1}(H_2^{-1}(1/2))$, there are nonadditive $((n, 2^k, d))$ quantum codes with rate $k/n \geq 1 - 2H_2(d/n)$.* \square

A Strongly Nonadditive Code. In this section we provide an example of a strongly nonadditive quantum error-correcting code. This is an $((11, 2, 3))$ strongly nonadditive code.

Consider the (Paley type) Hadamard matrix of order 12 (see, e.g., [10], p. 48). Delete the all-1 column and replace -1 by 1 and $+1$ by 0. The result is the following matrix

$$H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & \end{bmatrix}.$$

We denote the i^{th} row of H by r_i . The set $\mathcal{C} = \{r_i : 1 \leq i \leq 12\}$ is an $(11, 12, 6)$ code. Then a basis for the desired quantum code consists of the following two vectors:

$$\begin{aligned} |0_L\rangle &= \sum_{i=1}^{12} |r_i\rangle, \\ |1_L\rangle &= \sum_{i=1}^{12} |\mathbf{1} + r_i\rangle, \end{aligned}$$

where $\mathbf{1}$ is the all-1 vector of length 11. We claim these vectors are basis for an $((11, 2, 3))$ quantum code. We have to show that

$$\langle 0_L | X_\alpha Z_\beta | 0_L \rangle = 0, \quad (12)$$

$$\langle 1_L | X_\alpha Z_\beta | 1_L \rangle = 0, \quad (13)$$

$$\langle 0_L | X_\alpha Z_\beta | 1_L \rangle = 0, \quad (14)$$

for every $\alpha, \beta \in \{0, 1\}^{11}$ such that $1 \leq \text{wt}(\alpha \cup \beta) \leq 2$. First note that the distance of any two distinct vectors in the set

$$\{r_i : 1 \leq i \leq 12\} \cup \{\mathbf{1} + r_i : 1 \leq i \leq 12\}$$

is at least 5. Thus if $1 \leq \text{wt}(\alpha) \leq 4$ then all conditions (12)–(14) hold. Now suppose that $\alpha = \mathbf{0}$. Then (14) trivially holds. To see that (12) and (13) hold it is enough to note that if $1 \leq \text{wt}(\beta) \leq 2$ then $r_i \cdot \beta = 1$ for exactly 6 values of i . This completes the proof that $\{|0_L\rangle, |1_L\rangle\}$ is a basis for an $((11, 2, 3))$ quantum error-correcting code.

To show that this code is nonadditive, let $\varphi = (-1)^\lambda X_\alpha Z_\beta$ be any operator in the stabilizer of this code. Since $\varphi |0_L\rangle = |0_L\rangle$ and $\varphi |r_1\rangle = |\alpha\rangle$, hence $\lambda = 0$ and α should be one of r_i 's. Then we should have $\alpha = r_1 = \mathbf{0}$, because for every r_i , $i \neq 1$, there is some j such that $r_i + r_j$ is not equal to any r_k . Therefore, $\varphi = Z_\beta$. Then

$$Z_\beta |1_L\rangle = \sum_{i=1}^{12} (-1)^{(\mathbf{1}+r_i) \cdot \beta} |\mathbf{1} + r_i\rangle = \sum_{i=1}^{12} |\mathbf{1} + r_i\rangle$$

implies that $(\mathbf{1} + r_i) \cdot \beta = 0$, for every i . But the set $\{\mathbf{1} + r_i : 1 \leq i \leq 12\}$ has rank 11, so $\beta = \mathbf{0}$. This shows that the identity operator is the only operator in the stabilizer of this code. Finally, suppose that $X_\alpha T$ is in the generalized stabilizer of this code, where the operator T is of the form (7). Note that the operator T only affects the phases of the states, so the above argument also implies $\alpha = \mathbf{0}$. Now Theorem 5 implies that this code is strongly nonadditive.

4 Concluding Remarks

We showed that there are nonadditive codes with different minimum distances. We also showed that nonadditive codes that correct t errors can reach the asymptotic rate $R \geq 1 - 2H_2(2t/n)$. We introduced the notion of strongly nonadditive codes, and gave an example of such codes. It would be interesting to find more examples of such codes. We conjecture that the nonadditive codes constructed in Section 3.2 are also strongly nonadditive codes.

Recently we have improved the construction method for nonadditive quantum codes. With this new scheme, we are now able to give explicit constructions of nonadditive $((2m, \frac{1}{4}2^{2m}, 2))$ and strongly nonadditive $((2m + 1, \frac{1}{8}(1 - \frac{1}{2m})2^{2m+1}, 2))$ codes. Also we have improved the asymptotic Gilbert–Varshamov bound for the rate of nonadditive codes. The new bound, which is for *strongly* nonadditive codes, is the same as the bound for additive codes [4], i.e., $R \geq 1 - H_2(2t/n) - (2t/n) \log_2 3$. All these results will appear in the final version of this paper [14].

References

1. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin and W. K. Wootters, “Mixed state entanglement and quantum error correction,” *Phys. Rev. A*, Vol. 54, No. 5, pp. 3824–3851 (1996).
2. M. Grassl and Th. Beth, “A note on non-additive quantum codes,” LANL e-print quant-ph/97030126.
3. A. R. Calderbank, E. M. Rains, P. W. Shor and N. J. A. Sloane, “Quantum error correction and orthogonal geometry,” *Phys. Rev. Lett.*, Vol. 78, pp. 405–408 (1997).
4. A. R. Calderbank, E. M. Rains, P. W. Shor and N. J. A. Sloane, “Quantum error correction via codes over $\text{GF}(4)$,” LANL e-print quant-ph/9608006.
5. A. R. Calderbank and P. W. Shor, “Good quantum error-correcting codes exist,” *Phys. Rev. A*, Vol. 54, No. 2, pp. 1098–1105 (1996).
6. R. Cleve, “Quantum stabilizer codes and classical linear codes,” LANL e-print quant-ph/9612048.
7. D. Gottesman, “A class of quantum error-correcting codes saturating the quantum Hamming bound,” *Phys. Rev. A*, Vol. 54, No. 3, pp. 1862–1868 (1996).
8. E. Knill and R. Laflamme, “A theory of quantum error-correcting codes,” LANL e-print quant-ph/9604034.
9. F. J. MacWilliams, N. J. Sloane and J. P. Thompson, “Good self dual codes exist,” *Discrete Math.*, vol. 3, pp. 153–162 (1972).
10. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*, North-Holland, New York, 1977.

11. E. M. Rains, "Quantum shadow enumerators," LANL e-print quant-ph/9611001.
12. E. M. Rains, "Quantum codes of minimum distance two," LANL e-print quant-ph/9704043.
13. E. M. Rains, R. H. Hardin, P. Shor and N. J. A. Sloane, "A nonadditive quantum code," *Phys. Rev. Lett.*, Vol. 79, pp. 953–954 (1997).
14. V. Roychowdhury and F. Vatan, "On the structure of additive codes and the existence of nonadditive codes," LANL e-print quant-ph/9710031.
15. A. M. Steane, "Error correcting codes in quantum theory," *Phys. Rev. Lett.*, Vol. 77, No. 5, pp. 793–797 (1996).
16. F. Vatan, V. P. Roychowdhury and M. P. Anantram, "Spatially correlated qubit errors and burst-correcting quantum codes," LANL e-print quant-ph/9704019. To appear in *IEEE Transactions on Information Theory*.

Quantum Error Correction Is Applicable for Reducing Spatially Correlated Decoherence

Lu-Ming Duan and Guang-Can Guo

Department of Physics and Nonlinear Science Center,
University of Science and Technology of China,
Hefei 230026, People's Republic of China
gcguo@sun1x06.nsc.ustc.edu.cn

Abstract. All the quantum error correction schemes devised to suppress independent decoherence are shown to be applicable for reducing general spatially correlated decoherence.

PACS numbers: 03.75, 03.65.Bz, 89.70.+c, 42.50.Dv

Key words: Quantum error correction, independent decoherence, spatially correlated decoherence, quantum trajectory theory.

The condition of independent decoherence plays an important role in devising the quantum error correcting codes (QECCs) [1-3]. However, in practice, spatial correlation may take place in decoherence of qubits [4-6]. Are the QECCs devised to correct single-qubit errors applicable for reducing spatially correlated decoherence? We calculate the state fidelity after applying the conventional quantum error correction procedure in the case of spatially correlated decoherence, and show that the QECCs are still valid. This extends the decoherence model in the previous quantum error correction schemes to more realistic circumstances.

We take the analysis of spatially correlated phase damping as an example to illustrate the general result. It is not difficult to derive that the master equation for spatially correlated phase damping has the form [7]

$$\frac{d}{dt}\rho(t) = \frac{\varepsilon}{2} \sum_{i,j=1}^L \left\{ \gamma_{ij} \left[2\sigma_i^z \rho(t) \sigma_j^z - \sigma_j^z \sigma_i^z \rho(t) - \rho(t) \sigma_j^z \sigma_i^z \right] \right\}, \quad (1)$$

where $\rho(t)$ is the reduced density of the qubits in the interaction picture, and σ_i^z are all Pauli's operators describing the qubits. The coefficients γ_{ij} , whose values depend on the correlations of the bath operators and the separations of the qubits, completely characterize the spatial correlation properties. L is the number of qubits and ε is a normalization constant to make γ_{ij} satisfy the normalization condition $\sum_{i=1}^L \gamma_{ii} = 1$. Define an $L \times L$ spatial correlation matrix

Γ by $\Gamma = [\gamma_{ij}]$. Γ is a real symmetric matrix, thus it can be diagonalized by an orthogonal matrix O as $O\Gamma O^+ = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_L)$, where $\gamma_1, \gamma_2, \dots$, and γ_L , satisfying $\sum_{i=1}^L \gamma_i = 1$, are the eigenvalues of the matrix Γ . Define the operators s_l^z by the equation $(s_1^z, s_2^z, \dots, s_L^z) = (\sigma_1^z, \sigma_2^z, \dots, \sigma_L^z) O^+$. The master equation (3) is then rewritten as

$$\frac{d}{dt}\rho(t) = -iH_{eff}\rho(t) + i\rho(t)H_{eff}^\dagger + \sum_{l=1}^L \varepsilon\gamma_l s_l^z \rho(t) s_l^z, \quad (2)$$

where the non-Hermitian effective Hamiltonian is

$$H_{eff} = -\frac{i\varepsilon}{2} \sum_{l=1}^L \gamma_l (s_l^z)^2. \quad (3)$$

In the language of quantum trajectories [8,9], the system evolution described by Eq. (2) is represented by an ensemble of wave functions that propagate according to the effective Hamiltonian H_{eff} , interrupted at random times by quantum jumps. In a finite time T_0 , we perform N times error correction. The state of the qubits, initially denoted by $|\Psi(0)\rangle$, evolves in the time interval $\Delta t = T_0/N$ according to Eq. (2). Suppose that $\epsilon = \max(\varepsilon\gamma_l\Delta t) \ll 1$. Up to order ϵ , the normalized state of the qubits after Δt will be either

$$|\Psi_l(\Delta t)\rangle = \sqrt{\frac{\varepsilon\gamma_l\Delta t}{p_l}} s_l^z |\Psi(0)\rangle \quad (4)$$

with probability $p_l = \varepsilon\gamma_l\Delta t \langle\Psi(0)| (s_l^z)^2 |\Psi(0)\rangle$ in case of a jump in decay channel l at a random time in the interval Δt , or

$$\begin{aligned} |\Psi_0(\Delta t)\rangle &= \frac{1}{\sqrt{p_0}} \exp(-iH_{eff}\Delta t) |\Psi(0)\rangle \\ &\approx \frac{1}{\sqrt{p_0}} \left\{ |\Psi(0)\rangle - \frac{\varepsilon\Delta t}{2} \sum_{l=1}^L \left[\gamma_l (s_l^z)^2 \right] |\Psi(0)\rangle \right\} \end{aligned} \quad (5)$$

with probability $p_0 \approx 1 - \varepsilon\Delta t \langle\Psi(0)| \sum_{l=1}^L \left[\gamma_l (s_l^z)^2 \right] |\Psi(0)\rangle - o(\epsilon^2)$ if no jump occurred. Let $s_0^z = \sum_{l=1}^L \left[\gamma_l (s_l^z)^2 \right]$, which represents the first order error due to the effective evolution, whereas s_l^z ($l = 1, 2, \dots, L$) can be explained as the first order errors due to quantum jumps. For independent decoherence, s_0^z and s_l^z reduce to the single qubit operators I and σ_l^z , respectively. But for general spatially correlated decoherence, the errors s_0^z and s_l^z do not have this desired property. Nevertheless, here we show, that the QECCs devised to correct single-qubit errors are still valid in general circumstances, and moreover, the encoding, decoding, and error correction operations remain the same as those in the case

of independent decoherence. In the QECCs, phase errors are represented by the operators σ_l^z ($l = 1, 2, \dots, L$). For simplicity of symbols, let $\sigma_0^z = I$, where I is the unit operator, corresponding that there is no error. We consider orthogonal codes [10]. Most of the discovered codes belong to this class. For any orthogonal phase error correcting codes, the encoded input state $|\Psi(0)\rangle$ satisfies [10]

$$\langle \Psi(0) | \sigma_i^z \sigma_j^z | \Psi(0) \rangle = \delta_{ij} \quad (6)$$

where $(i, j = 0, 1, \dots, L)$.

During the error correction procedure, we first detect the error syndrome. If there is a σ_i^z ($i = 0, 1, \dots, L$) error (a σ_0^z error in fact means that there is no error.), i.e., the state becomes $\sigma_i^z |\Psi(0)\rangle$, we apply the recovery operator σ_i^z to the state and thus get the initial encoded state $|\Psi(0)\rangle$. In the case of spatially correlated decoherence, phase errors are represented not by σ_i^z , but by s_i^z instead. However, we still adopt the above error correction procedure. If there is a s_i^z ($i = 0, 1, \dots, L$) error, which occurs with probability p_i , we detect the error syndrome, and with probability $|\langle \Psi(0) | \sigma_j^z | \Psi_i(\Delta t) \rangle|^2$ find that the error is σ_j^z . After this detection, the state is collapsed into $\sigma_j^z |\Psi(0)\rangle$. Then we apply the recovery operator σ_j^z and get the initial state $|\Psi(0)\rangle$. The whole error correction procedure described above yields the following average state fidelity

$$\begin{aligned} F_c(\Delta t) &= \sum_{i=0}^L \sum_{j=0}^L p_i |\langle \Psi(0) | \sigma_j^z | \Psi_i(\Delta t) \rangle|^2 \\ &= 1 - o(\epsilon^2). \end{aligned} \quad (7)$$

In deriving Eq. (7), we used Eq. (6) and the identity $\sum_{j=1}^L O_{ij}^2 = 1$. After the whole time T_0 , the final average fidelity $F_c(T_0)$ is approximately

$$F_c(T_0) \approx [F_c(\Delta t)]^N \approx 1 - o(N\epsilon^2). \quad (8)$$

Since $\epsilon = \max(\varepsilon \gamma_l \Delta t) \propto \frac{1}{N}$, $N\epsilon^2$ can be made arbitrarily small by a frequent repetition of the error correction procedure. This completes the proof that the conventional QECCs are valid for reducing spatially correlated decoherence.

Acknowledgment

This project was supported by the National Nature Science Foundation of China.

References

1. P. W. Shor, Phys. Rev. A 52, R2493 (1995);
2. T. Pellizzari, Th. Beth, M. Grassl, and J. Muller-Quade, Phys. Rev. A 54, 2698 (1996).
3. A. M. Steane, LANL e-print quant-ph/9708022.

4. G. M. Palma, K. A. Suominen, and A. K. Ekert, Proc. R. Soc. London A 452, 567 (1996).
5. L. M. Duan and G. C. Guo, Phys. Rev. Lett. 79, 1953 (1997); Phys. Rev. A 56, 4466 (1997); *ibid* 57, 737 (1998).
6. P. Zanardi, LANL e-print quant-ph/9705045.
7. C. W. Gardiner, *Quantum Noise* (Springer-Verlag, Berlin, 1991).
8. C. W. Gardiner, A. S. Parkins, and P. Zoller, Phys. Rev. A 46, 4363 (1992).
9. R. Dum, A. S. Parkins, P. Zoller, and C. W. Gardiner, Phys. Rev. A 46, 4382 (1992).
10. E. Knill and R. Laflamme, Phys. Rev. A 55, 900 (1997).

Topological Quantum Computation

R. Walter Ogburn and John Preskill

California Institute of Technology, Pasadena, CA 91125, USA
reuben@cco.caltech.edu, preskill@theory.caltech.edu

Abstract. Following a suggestion of A. Kitaev, we explore the connection between fault-tolerant quantum computation and nonabelian quantum statistics in two spatial dimensions. A suitably designed spin system can support localized excitations (quasiparticles) that exhibit long-range nonabelian Aharonov-Bohm interactions. Quantum information encoded in the charges of the quasiparticles is highly resistant to decoherence, and can be reliably processed by carrying one quasiparticle around another. If information is encoded in pairs of quasiparticles, then the Aharonov-Bohm interactions can be adequate for universal fault-tolerant quantum computation.

1 Fault-Tolerant Quantum Computation

Quantum computers appear to be capable, at least in principle, of solving certain problems far faster than any conceivable classical computer [1]-[3]. In practice, though, quantum computing technology is still in its infancy. While a practical and useful quantum computer may eventually be constructed, we cannot clearly envision at present what the hardware of that machine will be like. Nevertheless, we can be quite confident that any practical quantum computer will incorporate some type of error correction into its operation. Quantum computers are far more susceptible to making errors than conventional digital computers [4]-[8], and some method of controlling and correcting those errors will be needed to prevent a quantum computer from crashing.

The future prospects for quantum computing received a tremendous boost from the discovery by Peter Shor [9] and Andrew Steane [10,11] that quantum error correction is really possible in principle. But this discovery in itself is not sufficient to ensure that a noisy quantum computer can perform reliably. To carry out a quantum error-correction protocol, we must first encode the quantum information we want to protect, and then repeatedly perform recovery operations that reverse the errors that accumulate. Since encoding and recovery are themselves complex quantum computations, errors will inevitably occur while we perform these operations. Thus, we need to find methods for recovering from errors that are sufficiently robust to succeed with high reliability even when we make some errors during the recovery step. Such fault-tolerant recovery methods were first developed by Shor [12] and Alexei Kitaev [13]; these methods were later generalized and improved by Shor and David DiVincenzo [14], and by Steane [15].

Furthermore, to operate a quantum computer, we must do more than just *store* quantum information; we must *process* the information. We need to be able to perform quantum gates, in which two or more encoded qubits come together and interact with one another. If an error occurs in one qubit, and then that qubit interacts with another through the operation of a quantum gate, the error is likely to spread to the second qubit. We must design our gates to minimize the propagation of error. The central challenge is to construct a universal set of quantum gates that can act on the encoded data blocks without introducing an excessive number of errors. Such a scheme for fault-tolerant quantum computation was first developed by Shor [12] and later generalized by Daniel Gottesman [16].

Once the elementary gates of our quantum computer are sufficiently reliable, we can perform fault-tolerant quantum gates on encoded information, along with fault-tolerant error recovery, to improve the reliability of the device. But for any fixed quantum code, or even for most infinite classes that contain codes of arbitrarily large block size, these procedures will eventually fail if we attempt a very long computation. However, there is a special class of codes (*concatenated codes*) which enable us to perform longer and longer quantum computations reliably, as we increase the block size at a modest rate [17]–[23]. Invoking concatenated codes we can establish an *accuracy threshold* for quantum computation; once our hardware meets a specified standard of accuracy, quantum error-correcting codes and fault-tolerant procedures enable us to perform arbitrarily long quantum computations with arbitrarily high reliability.

With the development of fault-tolerant methods, we now know that it is possible in principle for the operator of a quantum computer to actively intervene to stabilize the device against errors in a noisy (but not *too* noisy) environment. In the long term, though, fault tolerance might be achieved in practical quantum computers by a rather different route—with intrinsically fault-tolerant hardware. Such hardware, designed to be impervious to *localized* influences, could be operated relatively carelessly, yet could still store and process quantum information robustly.

In this paper, we explore a scheme for fault-tolerant hardware envisioned by Kitaev [24], in which the quantum gates exploit nonabelian Aharonov-Bohm interactions among distantly separated quasiparticles in a suitably constructed two-dimensional spin system. Though the laboratory implementation of Kitaev’s idea may be far in the future, his work offers a new slant on quantum fault tolerance that shuns the analysis of abstract quantum circuits, in favor of new physics principles that might be exploited in the reliable processing of quantum information.

We explain in §2 that charges participating in long-range Aharonov-Bohm phenomena are impervious to local disturbances, so that quantum information encoded in such charges is robust. In §3 we argue that nonabelian Aharonov-Bohm interactions among quasiparticles arise in a class of two-dimensional spin systems. These interactions are discussed in detail in §4; we see that the exchange of two quasiparticles can modify the charges carried by the particles; thus par-

ticles with *different* charges may actually be *indistinguishable*. In particular, a quasiparticle that carries a superposition of two different charges need not decohere, because the local environment is indifferent to the value of the charge. In §5 we sketch our main result: that nonabelian Aharonov-Bohm interactions are adequate for universal quantum computation, in a model with a sufficiently rich group-theoretic structure. We conclude in §6 with some tentative speculations regarding the implications of quantum fault tolerance for fundamental physics.

Recent claims about the potential for the fault-tolerant manipulation of complex quantum states may seem grandiose from the perspective of present-day technology. Surely, we have far to go before devices are constructed that can, say, exploit the accuracy threshold for quantum computation [25]. Nevertheless, we feel strongly that recent work relating to quantum error correction will have an enduring legacy. Theoretical quantum computation has developed at a spectacular pace over the past three years. If, as appears to be the case, the quantum classification of computational complexity differs from the classical classification, then no conceivable classical computer can accurately predict the behavior of even a modest number of qubits (of order 100). Perhaps, then, relatively small quantum systems will have far greater potential than we now suspect to surprise, baffle, and delight us. Yet this potential could never be realized were we unable to protect such systems from the destructive effects of noise and decoherence. Thus the discovery of fault-tolerant methods for quantum error recovery and quantum computation has exceptionally deep implications, both for the future of experimental physics and for the future of technology. The theoretical advances have illuminated the path toward a future in which intricate quantum systems may be persuaded to do our bidding.

2 Aharonov-Bohm Phenomena and Superselection Rules

Topological concepts have a natural place in the discussion of quantum error correction and fault-tolerant computation. Topology concerns the “global” properties of an object that remain unchanged when we deform the object locally. The central idea of quantum error correction is to store and manipulate quantum information in a “global” form that is resistant to local disturbances. A fault-tolerant gate should be designed to act on this global information, so that the action it performs on the encoded data remains unchanged even if we deform the gate slightly; that is, even if the implementation of the gate is not perfect.

In seeking physical implementations of fault-tolerant quantum computation, then, we ask whether there are known systems in which physical interactions have a topological character. Indeed, topology is at the essence of the *Aharonov-Bohm effect*. If an electron is transported around a perfectly shielded magnetic solenoid, its wave function acquires a phase $e^{ie\Phi}$, where e is the electron charge and Φ is the magnetic flux enclosed by the solenoid. This Aharonov-Bohm phase is a topological property of the path traversed by the electron — it depends only on how many times the electron circumnavigates the solenoid, and is unchanged when the path is smoothly deformed. (See Fig. [11](#)) We are thus led to contemplate

a realization of quantum computation in which information is encoded in a form that can be measured and manipulated through Aharonov-Bohm interactions — topological interactions that are immune to local disturbances.

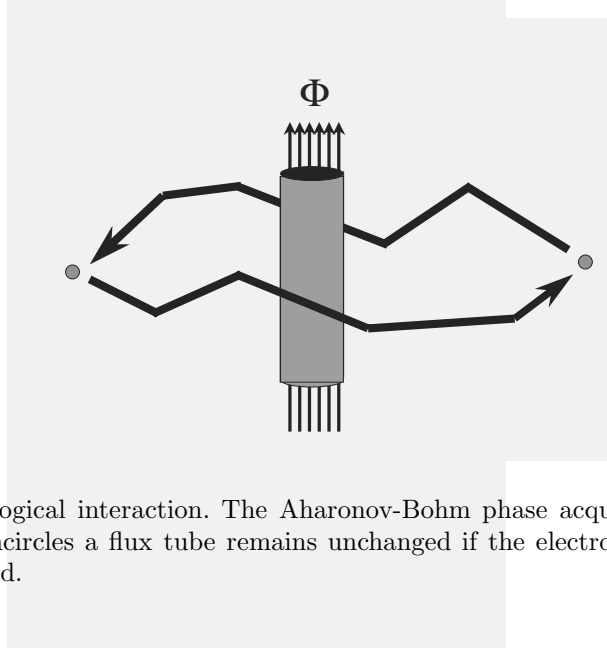


Fig. 1. A topological interaction. The Aharonov-Bohm phase acquired by an electron that encircles a flux tube remains unchanged if the electron's path is slightly deformed.

It is useful to reexpress this reasoning in the language of superselection rules. A superselection rule, as we are using the term here, arises (in a field theory or spin system defined in an infinite spatial volume) if Hilbert space decomposes into mutually orthogonal sectors, where each sector is preserved by any local operation. Perhaps the most familiar example is the charge superselection rule in quantum electrodynamics. An electric charge has an infinite range electric field. Therefore no local action can create or destroy a charge, for to destroy a charge we must also destroy the electric field lines extending to infinity, and no local procedure can accomplish this task.

The Aharonov-Bohm interaction is also an infinite range effect; the electron acquires an Aharonov-Bohm phase upon circling the solenoid no matter what its distance from the solenoid. So we may infer that no local operation can destroy a charge that participates in Aharonov-Bohm phenomena. If we consider two objects carrying such charges, widely separated and well isolated from other charged objects, then any process that changes the charge on either object would have to act coherently in the whole region containing the two charges. Thus, the charges are quite robust in the presence of localized disturbances; we can strike the particle with a hammer or otherwise abuse it without modifying the charges that it carries.

Following Kitaev [24], we may envision a *topological quantum computer*, a device in which quantum information is encoded in the quantum numbers carried by quasiparticles that reside on a two-dimensional surface and have long-range

Aharonov-Bohm interactions with one another. At zero temperature, an accidental exchange of quantum numbers between quasiparticles (an error) arises only due to quantum tunneling phenomena involving the virtual exchange of charged objects. The amplitude for such processes is of the order of e^{-mL} , where m is the mass of the lightest charged object (in natural units), and L is the distance between the two quasiparticles. If the quasiparticles are kept far apart, the probability of an error afflicting the encoded information will be extremely low. At finite temperature T , there is an additional source of error, because an uncontrolled plasma of charged particles will inevitably be present, with a density proportional to the Boltzmann factor $e^{-\Delta/T}$, where Δ is the mass gap (not necessarily equal to the “curvature mass” m). Sometimes one of the plasma particles will slip unnoticed between two of our data-carrying particles, resulting in an exchange of charge and hence an error. To achieve an acceptably low error rate, then, we would need to keep the temperature well below the gap Δ (or else we would have to monitor the thermal plasma very faithfully).

3 The Fractional Quantum Hall Effect and Beyond

If our device is to be capable of performing interesting computations, the Aharonov-Bohm phenomena that it employs must be *nonabelian*. Only then will we be able to build up complex unitary transformations by performing many particle exchanges in succession. Such nonabelian Aharonov-Bohm effects can arise in systems with nonabelian gauge fields. Nature has been kind enough to provide us with some fundamental nonabelian gauge fields, but unfortunately not very many, and none of these seem to be suited for practical quantum computation. To realize Kitaev’s vision, then, we must hope that nonabelian Aharonov-Bohm effects can arise as complex collective phenomena in (two-dimensional electron or spin) systems that have only short-range fundamental interactions.

In fact, one of the most remarkable discoveries of recent decades has been that infinite range Aharonov-Bohm phenomena *can* arise in such systems, as revealed by the observation of the fractional quantum Hall effect. The electrons in quantum Hall systems are so highly frustrated that the ground state is an extremely entangled state with strong quantum correlations extending out over large distances. Hence, when one quasiparticle is transported around another, even when the quasiparticles are widely separated, the many electron wave function acquires a nontrivial Berry phase (such as $e^{2\pi i/3}$). This Berry phase is indistinguishable in all its observable effects from an Aharonov-Bohm phase arising from a fundamental gauge field, and its experimental consequences are spectacular [26].

The Berry phases observed in quantum Hall systems are abelian (although there are some strong indications that nonabelian Berry phases can occur under the right conditions [27, 28]), and so are not very interesting from the viewpoint of quantum computation. But Kitaev [24] has described a family of simple spin systems with local interactions in which the existence of quasiparticles with nonabelian Berry phases can be demonstrated. (The Hamiltonian of the system so frustrates the spins that the ground state is a highly entangled state with infinite

range quantum correlations.) These models are sufficiently simple (although unfortunately they require four-body interactions), that one can imagine a designer material that can be reasonably well-described by one of Kitaev’s models. The crucial topological properties of the model are relatively insensitive to the precise microscopic details, so the task of the fabricator who “trims” the material may not be overly demanding. If furthermore it were possible to control the transport of individual quasiparticles (perhaps with a suitable magnetic tweezers), then the system could be operated as a fault-tolerant quantum computer.

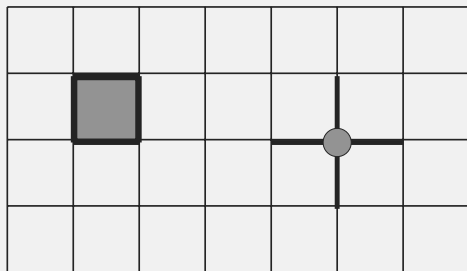


Fig. 2. A Kitaev spin model. Spins reside on the lattice links. The four spins that meet at a site or share a plaquette are coupled.

To construct his models, Kitaev considers a square lattice, with spins residing on each lattice link. The Hamiltonian is expressed as a sum of mutually commuting four-body operators, one for each site and one for each plaquette of the lattice. (See Fig. 2.) Because the terms are mutually commuting, it is simple to diagonalize the Hamiltonian by diagonalizing each term separately. The operators on sites resemble local gauge symmetries (acting independently at each site), and a state that minimizes these terms is invariant under the local symmetry, like the physical states that obey Gauss’s law in a gauge theory. The operators on plaquettes are like “magnetic flux” operators in a gauge theory, and these terms are minimized when the magnetic flux vanishes everywhere. The excitation spectrum includes states in which Gauss’s law is violated at isolated sites — these points are “electrically charged” quasiparticles — and states in which the magnetic flux is nonvanishing at isolated plaquettes — these are magnetic fluxon quasiparticles. The quantum entanglement of the ground state is such that a nontrivial Berry phase is associated with the transport of a charge around a flux — this phase is identical to the Aharonov-Bohm phase in the analog gauge theory.

These Aharonov-Bohm phenomena are stable even as we deform the Hamiltonian of the theory. Indeed, if the deformation is sufficiently small, we can study its effects using perturbation theory. But as long as the perturbations are local in

space, topological effects are robust, since perturbation theory is just a sum over localized influences. Whatever destroys the long-range topological interactions must be nonperturbative in the deformation of the theory.

Two types of nonperturbative effects can be anticipated[29]. The ground state of the theory might become a “flux condensate” with an indefinite number of magnetic excitations. In this event, there would be a long-range attractive interaction between charged particles and their antiparticles. It would be impossible to separate charges, and there would be no long-range effects. In a gauge theory, this phenomenon would be called *electric confinement*. Alternatively, a condensate of electric quasiparticles might appear in the ground state. Then the magnetic excitations would be confined, and again the long-range Aharonov-Bohm effects would be destroyed. In a gauge theory, we would call this the Higgs phenomenon (or magnetic confinement).

Thus, as we deform Kitaev’s Hamiltonian, we can anticipate that a phase boundary will eventually be encountered, beyond which either electric confinement or the Higgs phenomenon will occur. The size of the region enclosed by this boundary will determine how precisely a material will need to be fabricated in order to behave as Kitaev specifies. A particularly urgent question for the material designer is whether cleverly chosen *two-body* interactions might so frustrate a spin system as to produce a highly entangled ground state and nonabelian Aharonov-Bohm interactions among the quasiparticle excitations.

The fractional quantum Hall effect, and Kitaev’s models, speak a memorable lesson. We find gauge phenomena emerging as collective effects in systems with only short range interactions. It is intriguing to speculate that the gauge symmetries known in Nature could have a similar origin.

4 Topological Interactions

As we have noted, in Kitaev’s spin models, there are two types of charges that can be carried by localized quasiparticles, which we may call “electric” and “magnetic” charges. In the simplest type of model, the “magnetic flux” carried by a particle can be labeled by an element of a finite group G , and “electric charges” are labeled by irreducible representations¹ of G . If a charged particle in the irreducible representation $D^{(\nu)}$, whose quantum numbers are encoded in an internal wavefunction $|\psi^{(\nu)}\rangle$, is carried around a flux labeled by group element $u \in G$, then the wavefunction is modified according to

$$|\psi^{(\nu)}\rangle \rightarrow D^{(\nu)}(u)|\psi^{(\nu)}\rangle . \quad (1)$$

Exploiting this interaction, we can *measure* a magnetic flux by scattering a suitable charged particle off of the flux[30]. For example, we could construct a Mach-Zender flux interferometer as shown in Fig. 3 that is sensitive to the

¹ There can also be “dyons” that carry both types of charge, and the classification of the charge carried by a dyon is somewhat subtle, but we will not need to discuss explicitly the properties of the dyons.

relative phase acquired by the charged particle paths that pass to the left or right of the flux. If we balance the interferometer properly, we can distinguish between, say, two flux values $u_1, u_2 \in G$; a u_1 flux will be detected emerging from one arm of the interferometer, and a u_2 flux from the other arm. Of course, the interferometer we build will not be flawless, but the flux measurement can nevertheless be fault-tolerant — if we have many charged projectiles and perform the measurement repeatedly, we can determine the flux with very high statistical confidence.

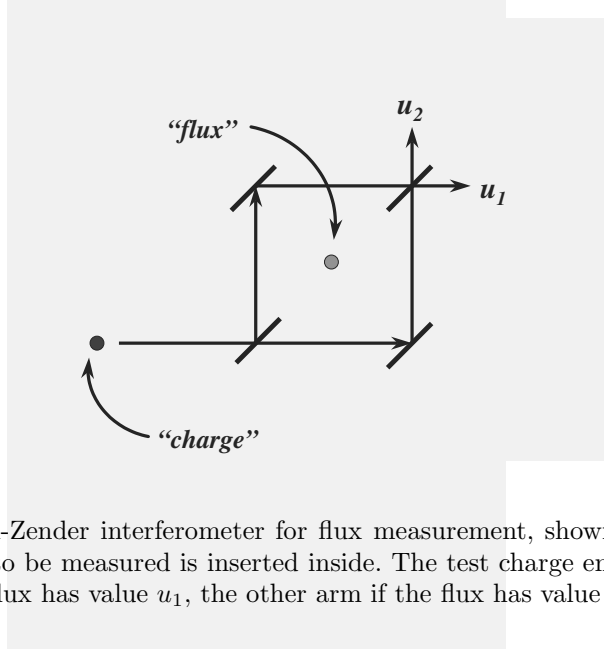


Fig. 3. A Mach-Zender interferometer for flux measurement, shown schematically. The flux to be measured is inserted inside. The test charge emerges from one arm if the flux has value u_1 , the other arm if the flux has value u_2 .

If the two fluxes u_1 and u_2 belong to the same conjugacy class in G , then there is a symmetry relating the two fluxons, so that all local physics is indifferent to the value of the flux (see below). Therefore, a coherent superposition of fluxes

$$a|u_1\rangle + b|u_2\rangle \quad (2)$$

will not readily decohere due to localized interactions with the environment. But the flux interferometer (operated repeatedly) will project the fluxon onto either of the flux eigenstates $|u_1\rangle$ (with probability $|a|^2$) or $|u_2\rangle$ (with probability $|b|^2$).

Now imagine that two fluxons have been carefully calibrated, so that one is known to carry the flux u_1 and the other the flux u_2 . And suppose that the two vortices are carefully “exchanged” by carrying the first around the second as shown in Fig. 4, and that we subsequently remeasure the fluxes. Carrying a charged particle around the fluxon on the right, after the exchange, is topologically equivalent to carrying the charged particle around first the right fluxon, then the left fluxon, and finally the right fluxon in the opposite direction, before the exchange. We infer that the exchange modifies the quantum numbers of the fluxons according to

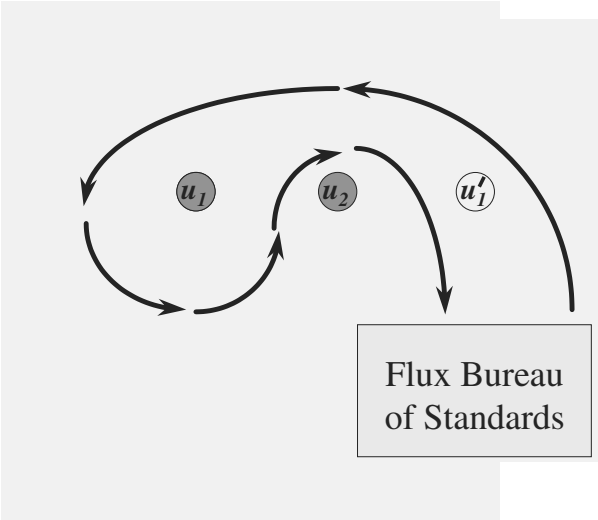


Fig. 4. The flux exchange interaction. The flux labeled u_1 is carried from its original position (shaded) to its new position (unshaded), and then remeasured. The charged particle path shown that encircles the original position of the flux is topologically equivalent to a path that encircles the new position; hence the value of the flux changes from u_1 to $u'_1 = u_2^{-1}u_1u_2$.

$$|u_1\rangle|u_2\rangle \rightarrow |u_2\rangle|u_2^{-1}u_1u_2\rangle, \quad (3)$$

a nontrivial interaction if the two fluxes fail to commute[31]. Thus, noncommuting fluxes have interesting Aharonov-Bohm interactions of their own, even in the absence of any electric charges. Because carrying one flux around another can *conjugate* the value of the flux, two fluxons carrying conjugate fluxes must be regarded as *indistinguishable* particles[32]. An exchange of two such objects can modify their internal quantum numbers; we will refer to them as *nonabelions*[33], indistinguishable particles in two dimensions that obey an exotic nonabelian variant of quantum statistics.

We will use the exchange interaction Eq. (3) as a fundamental logical operation in our Aharonov-Bohm quantum computer. However, it will actually be convenient to encode qubits in pairs of fluxons, where the total flux of the pair is trivial[24]. That is, we will consider fluxon-antifluxon pairs of the form $|u, u^{-1}\rangle$, but where the flux and antiflux are kept far enough apart from one another that an inadvertent exchange of quantum numbers between them is unlikely. To perform logic, we may pull one pair through another as shown in Fig. 5. Since the total flux that passes through the middle of the outside pair is trivial, this pair is not modified, but the inside fluxes are conjugated by the outside flux:

$$|u_1, u_1^{-1}\rangle|u_2, u_2^{-1}\rangle \rightarrow |u_2, u_2^{-1}\rangle|u_2^{-1}u_1u_2, u_2^{-1}u_1^{-1}u_2\rangle; \quad (4)$$

an operation that is evidently isomorphic to the effect of the exchange of single fluxes described by Eq. (3). Using pairs instead of single fluxons has two advantages. First, since each pair has trivial total flux, the pairs do not interact

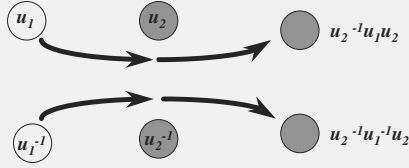


Fig. 5. The “pull-through” interaction. One flux pair is pulled through another. The outside flux is unmodified, but the inside flux is conjugated by the outside flux.

unless one is pulled through another; therefore, we can easily shunt pairs around the device without inducing any unwanted interactions with distant pairs. Second, and more important, pairs can carry charges even if each member of the pair carries no charge [34,35]. The charge of a pair can be measured, and this charge-measurement operation will be a crucial ingredient in the construction of a universal set of quantum gates. The operation Eq. (4) can be regarded as a *classical* logic gate; it takes flux eigenstates to flux eigenstates. To perform interesting quantum computations, we will need to be able to prepare coherent superpositions of flux eigenstates. This is what we can accomplish by measuring the charge of a pair.

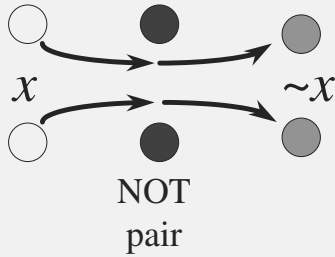


Fig. 6. The NOT gate. Pulling a computational flux pair through a NOT pair flips the value of the encoded bit.

Suppose that u_0 and $u_1 \in G$ are related by $u_1 = v^{-1}u_0v$ for some $v \in G$. Then if we think of the flux eigenstates $|u_0, u_0^{-1}\rangle$ and $|u_1, u_1^{-1}\rangle$ as computational basis states, the effect of pulling either pair through a $|v, v^{-1}\rangle$ pair can be interpreted as a NOT (or σ_x) gate:

$$|u_0, u_0^{-1}\rangle \leftrightarrow |u_1, u_1^{-1}\rangle \tag{5}$$

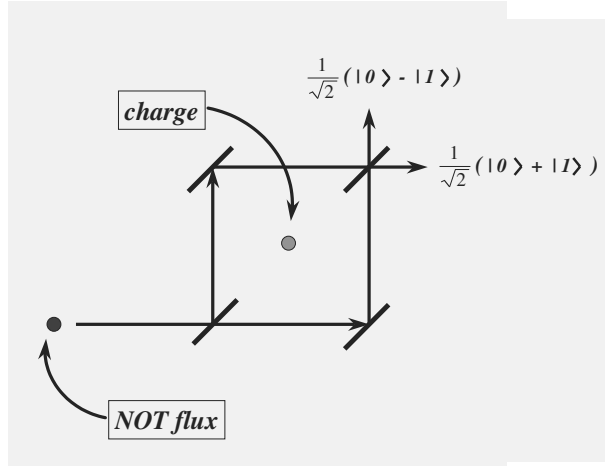


Fig. 7. A Mach-Zender interferometer for charge measurement, shown schematically. The flux pair whose charge is to be measured is inserted inside. If the test NOT flux emerges from one arm, the $|+\rangle$ charge state has been prepared; if it emerges from the other arm, $|-\rangle$ has been prepared.

(see Fig. 6). But suppose we wish to prepare one of the states

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|u_0, u_0^{-1}\rangle \pm |u_1, u_1^{-1}\rangle) . \quad (6)$$

We can project a coherent superposition of $|u_0, u_0^{-1}\rangle$ and $|u_1, u_1^{-1}\rangle$ onto the $\{|\pm\rangle\}$ basis by scattering a $|v\rangle$ fluxon off the pair, or in other words by operating a *charge interferometer*, as in Fig. 7. When the $|v\rangle$ fluxon navigates around the pair, it acquires a trivial Aharonov-Bohm phase if the pair is in the state $|+\rangle$ and the nontrivial phase -1 if the pair is in the state $|-\rangle$. If the interferometer is properly balanced, then, the $|v\rangle$ projectile will be detected emerging from one arm of the interferometer if the pair is $|+\rangle$, and the other arm if the pair is $|-\rangle$. This is an example of charge measurement. Though the interferometer will not be perfect, charge measurement (like flux measurement) can be fault-tolerant, if we repeat the measurement enough times.

5 Universal Topological Computation

Working with fluxon pairs as computational basis states, we have seen how to perform the exchange (or “pull through”) operation Eq. (4), how to measure flux (using previously calibrated charges), and how to measure charge (using previously calibrated fluxes). We will also suppose that we are able to produce a large supply of vortex pairs. Local processes produce pairs that carry no charge or flux; a charge-zero pair with trivial flux has the form (up to normalization)

$$|\text{charge zero}\rangle = \sum_u |u, u^{-1}\rangle , \quad (7)$$

where the sum ranges over a complete conjugacy class of G . Because this state is left invariant when conjugated by any element of G , it has trivial Aharonov-Bohm interactions with any flux, and so carries no detectable charge. After producing such a pair, we can perform flux measurement to project out one of the flux eigenstate pairs $|u, u^{-1}\rangle$. Performing many such measurements on many pairs, we can assemble a large reservoir of calibrated flux pairs that can be withdrawn as needed during the course of a computation.

But is our quantum computer universal — can we closely approximate any desired unitary transformation? To address this issue, we appeal to a theorem proved by Gottesman [16]. Suppose that we can perform any *classical* reversible operation; that is, any unitary transformation on n qubits that merely permutes the 2^n computational basis states. Then to achieve universal *quantum* computation, it is sufficient to be able to perform a few simple operations on individual qubits: the single-qubit gate σ_z , and measurement of the single-qubit observables σ_x , σ_y , and σ_z . In other words (if we envision the qubits as spin- $\frac{1}{2}$ objects), once we have a universal classical gate at our disposal, we can build a universal quantum computer if we are able to rotate a spin by 180° about the z axis [2] and can measure the spin along the x , y , and z axes.

In fact, there are groups G such that the operation Eq. (4) is sufficient for universal classical computation. The three-bit Toffoli gate, with action

$$\text{Toffoli} : |a, b, c\rangle \mapsto |a, b, c \oplus ab\rangle \quad (8)$$

on $a, b, c \in \{0, 1\}$, is a universal classical gate. We have found that a Toffoli gate can be constructed from Eq. (4) if $G = A_5$, the group of even permutations on five objects. We may, for example, choose computational basis states with

$$u_0 = (125) \ , \quad u_1 = (234) \ ; \quad (9)$$

that is, we choose our computational fluxes to be three-cycles with one object in common. Then a Toffoli gate can be constructed from a total of 16 elementary “pull-through” operations; six ancilla pairs are also used to catalyze this reaction. No Toffoli gate was found in any group smaller than A_5 [8]. Since A_5 is also the smallest of the finite nonsolvable groups, it is tempting to conjecture that nonsolvability is a necessary condition for universal classical computation generated by conjugation [4].

We have already remarked that an σ_x gate can be realized by pulling a computational vortex pair through the pair with flux v such that $u_1 = v^{-1}u_0v$; here we choose $v = (14)(35)$. It turns out that the σ_z gate can be constructed with

² Since σ_x is a classical gate, and $i\sigma_y = \sigma_z\sigma_x$, it follows that we can perform 180° rotations about each of the x , y and z axes.

³ Kitaev had reported earlier that universal classical computation is possible for $G = S_5$.

⁴ A finite group is *nonsolvable* if it has a nontrivial subgroup whose commutator subgroup is itself. Barrington [36] also found evidence for a separation in the computational complexity of group multiplication for solvable vs. nonsolvable groups.

six pull-through steps and four ancilla pairs. Measuring σ_z is the same as measuring flux, and we have already seen that σ_x measurement can be achieved by measuring the charge of a pair, specifically, by using a v projectile in a charge interferometer. It only remains to verify that we can measure σ_y . Though σ_y measurement cannot be carried out exactly in this scheme, it turns out that a *controlled- σ_y* gate can be constructed from 31 pull-through steps, and using 7 ancilla pairs. Appealing to another trick invented by Kitaev [37], we can use the controlled- σ_y gate repeatedly to carry out σ_y -measurement to any desired accuracy.⁵ Therefore, we have constructed a universal gate set using only the Aharonov-Bohm interactions of fluxes and charges; we have a fault-tolerant universal quantum computer.

Unfortunately, the spin model on which this construction is based is not so simple. Since the group A_5 has order 60, the Kitaev spin model that realizes this scenario has a 60-component spin residing at each lattice link (!) One hopes that a simpler implementation of universal Aharonov-Bohm computation will be found.

The fabrication of materials that emulate Kitaev's spin systems may lie far in the future. And even when such materials are available, there will be further challenges to the machine designer, such as finding a reliable way to shepherd individual quasiparticles along prescribed trajectories. In the nearer term, it is interesting to consider whether nontrivial quantum information processing might be feasible in existing quantum Hall systems. Furthermore, even if we are unable to operate an actual spin system as a quantum computer, a quantum cellular automaton that simulates the spin system may provide a promising paradigm for fault-tolerant quantum computation.

6 Is Nature Fault Tolerant?

The discovery of quantum error correction and fault tolerance has so altered our thinking about quantum information that it is appropriate to wonder about the potential implications for fundamental physics. And in fact, a fundamental issue pertaining to loss of quantum information has puzzled the physics community for over twenty years.

In 1975, Stephen Hawking [38] argued that quantum information is unavoidably lost when a black hole forms and then subsequently evaporates completely. The essence of the argument is very simple: because of the highly distorted causal structure of the black hole spacetime, the emitted radiation is actually on the *same* time slice as the collapsing body that disappeared behind the event horizon. If the quantum information that is initially encoded in the collapsing body is eventually to re-emerge encoded in the microstate of the emitted information, then that information must be in two places at once. In other words, the quantum information must be *cloned*, a known impossibility under the usual assumptions of quantum theory [39,40]. Hawking concludes that not all physical

⁵ Actually, what we really construct is a controlled- $\omega(i\sigma_y)$ gate where $\omega = e^{2\pi i/3}$, which is also adequate for measurement of σ_y .

processes can be governed by unitary time evolution; the laws of quantum theory need revision.

This argument is persuasive, but many physicists are very distrustful of the conclusion. Perhaps one reason for the skepticism is that it seems odd for Nature to tolerate just a little bit of information loss^[41]. If processes involving black holes can destroy information, then one expects that information loss is unsuppressed at the Planck length scale $(G\hbar/c^3)^{1/2} \sim 10^{-33}$ cm, a scale where virtual black holes continually arise as quantum fluctuations. It becomes hard to understand why quantum information can be so readily destroyed at the Planck scale, yet is so well preserved at the much longer distance scales that we have been able to explore experimentally — violations of quantum mechanics, after all, have never been observed.

Our newly acquired understanding of fault-tolerant quantum computation provides us with a fresh and potentially fruitful way to think about this problem. In Kitaev's spin models, we might imagine that localized processes that destroy quantum information are quite common. Yet were we to follow the evolution of the system with coarser resolution, tracking only the information encoded in the charges of distantly separated quasiparticles, we would observe unitary evolution to remarkable accuracy; we would detect no glimmer of the turmoil beneath the surface.⁶

Likewise, it is tempting to speculate that Nature has woven fault tolerance into her design, shielding the quantum noise at the Planck scale from our view. The discovery that quantum systems can be stabilized through suitable coding methods prompts us to ask the question: Is Nature fault tolerant? If so, then quantum mechanics may reign (to excellent accuracy) at intermediate length scales, but falter both at the Planck scale (where "errors" are common) and at macroscopic scales (where decoherence is rapid).

Acknowledgments

This work has been supported in part by DARPA under Grant No. DAAH04-96-1-0386 administered by the Army Research Office, by the Department of Energy under Grant No. DE-FG03-92-ER40701, and by Caltech's Summer Undergraduate Research Fellowship program. We are grateful for helpful conversations with David DiVincenzo, Daniel Gottesman, Michael Nielsen, and especially Alesha Kitaev.

References

1. R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).

⁶ Similar language could be used to characterize the performance of a concatenated code—errors are rare when we inspect the encoded information with poor resolution, but are seen to be much more common if we probe the code block at lower levels of concatenation.

2. D. Deutsch, Quantum theory, the Church-Turing principle and the universal quantum computer, *Proc. Roy. Soc. Lond. A* **400**, 96 (1985).
3. P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Fundamentals of Computer Science* (Los Alamitos, CA, IEEE Press, 1994), pp. 124-134.
4. R. Landauer, Is quantum mechanics useful? *Phil. Tran. R. Soc. Lond.* **353**, 367 (1995).
5. R. Landauer, The physical nature of information, *Phys. Lett. A* **217**, 188 (1996).
6. R. Landauer, Is quantum mechanically coherent computation useful? In *Proc. Drexel-4 Symposium on Quantum Nonintegrability-Quantum-Classical Correspondence*, Philadelphia, PA, 8 September 1994, ed. D. H. Feng and B.-L. Hu (Boston, International Press, 1997).
7. W. G. Unruh, Maintaining coherence in quantum computers, *Phys. Rev. A* **51**, 992 (1995).
8. S. Haroche and J. M. Raimond, Quantum computing: dream or nightmare? *Phys. Today* **49** (8), 51 (1996).
9. P. Shor, Scheme for reducing decoherence in quantum memory, *Phys. Rev. A* **52**, 2493 (1995).
10. A. M. Steane, Error correcting codes in quantum theory, *Phys. Rev. Lett.* **77**, 793 (1996).
11. A. M. Steane, Multiple particle interference and quantum error correction, *Proc. Roy. Soc. Lond. A* **452**, 2551 (1996).
12. P. Shor, Fault-tolerant quantum computation, in *Proceedings of the Symposium on the Foundations of Computer Science* (Los Alamitos, CA: IEEE Press, online preprint quant-ph/9605011, 1996).
13. A. Yu. Kitaev, Quantum error correction with imperfect gates, in *Quantum Communication, Computing and Measurement* ed. O. Hirota, A. S. Holevo, and C. M. Caves (New York, Plenum, 1997).
14. D. DiVincenzo and P. Shor, Fault-tolerant error correction with efficient quantum codes. *Phys. Rev. Lett.* **77**, 3260 (1996).
15. A. M. Steane, Active stabilization, quantum computation and quantum state synthesis, *Phys. Rev. Lett.* **78**, 2252 (1997).
16. D. Gottesman, A theory of fault-tolerant quantum computation, *Phys. Rev. A* (online preprint quant-ph/9702029, 1997).
17. E. Knill and R. Laflamme, Concatenated quantum codes (online preprint quant-ph/9608012, 1996).
18. E. Knill, R. Laflamme, and W. Zurek, Accuracy threshold for quantum computation, (online preprint quant-ph/9610011, 1996).
19. E. Knill, R. Laflamme, and W. Zurek, Resilient quantum computation: error models and thresholds *Science* **279**, 342 (1998).
20. D. Aharonov and M. Ben-Or, Fault tolerant quantum computation with constant error (online preprint quant-ph/9611025, 1996).
21. A. Yu. Kitaev, Quantum computing: algorithms and error correction, *Russian Math. Surveys* **6** (1997)..
22. J. Preskill, Reliable quantum computers, *Proc. R. Soc. Lond. A* **454**, 385 (1998).
23. C. Zalka, Threshold estimate for fault tolerant quantum computing (online preprint quant-ph/9612028, 1996).
24. A. Yu. Kitaev, Fault-tolerant quantum computation by anyons (online preprint quant-ph/9707021, 1997).
25. J. Preskill, Quantum computing: pro and con, *Proc. Roy. Soc. Lond. A* **454**, 469 (1998).

26. R. Prange and S. Girvin, eds., *The Quantum Hall Effect* (New York, Springer-Verlag, 1987).
27. N. Read and E. Rezayi, Quasiholes and fermionic zero modes of paired fraction quantum Hall states: the mechanism for nonabelian statistics (online preprint cond-mat/9609079, 1996).
28. C. Nayak and F. Wilczek, $2n$ quasihole states realize 2^{n-1} -dimensional spinor braiding statistics in paired quantum Hall states (online preprint cond-mat/9605145, 1996).
29. G. 't Hooft, On the phase transition toward permanent quark confinement, *Nucl. Phys. B* **138**, 1 (1978).
30. M. Alford, S. Coleman, and J. March-Russell, Disentangling nonabelian discrete quantum hair, *Nucl. Phys. B* **351**, 735 (1991).
31. F. A. Bais, Flux metamorphosis, *Nucl. Phys. B* **170**, 32 (1980).
32. H.-K. Lo and J. Preskill, Nonabelian vortices and nonabelian statistics, *Phys. Rev. D* **48**, 4821 (1993).
33. G. Moore and N. Read, Nonabelions in the fractional quantum Hall effect, *Nucl. Phys. B* **360**, 362 (1991).
34. M. G. Alford, K. Benson, S. Coleman, J. March-Russell, and F. Wilczek, Interactions and excitations of nonabelian vortices, *Phys. Rev. Lett.* **64**, 1632 (1990).
35. J. Preskill and L. M. Krauss, Local discrete symmetry and quantum mechanical hair, *Nucl. Phys. B* **341**, 50 (1990).
36. D. A. Barrington, Bounded width polynomial size branching programs recognize exactly those languages in NC^1 , *J. Comp. Sys. Sci.* **38**, 150-164 (1989).
37. A. Yu. Kitaev, Quantum measurements and the abelian stabilizer problem (online preprint quant-ph/9511026, 1995).
38. S. W. Hawking, Breakdown of predictability in gravitational collapse, *Phys. Rev. D* **14**, 2460 (1976).
39. D. Dieks, Communication by electron-paramagnetic-resonance devices. *Phys. Lett. A* **92**, 271 (1982).
40. W. K. Wootters, and W. H. Zurek, A single quantum cannot be cloned, *Nature* **299**, 802 (1982).
41. T. Banks, M. E. Peskin, and L. Susskind, Difficulties for the evolution of pure states into mixed states, *Nucl. Phys. B* **244**, 125 (1984).

NMR GHZ

R. Laflamme¹, E. Knill², W.H. Zurek¹, P. Catasti³, and S.V.S. Mariappan³

¹ Theoretical Astrophysics T-6, MS B-288;

² Computer Research and Applications CIC-3, MS B-265;

³ Theoretical Biology T-10, MS B-288

Los Alamos National Laboratory, Los Alamos, NM 87455

Abstract. We describe the creation of a Greenberger-Horne-Zeilinger (GHZ) state of the form $(|000\rangle + |111\rangle)/\sqrt{2}$ (three maximally entangled quantum bits) using Nuclear Magnetic Resonance (NMR). We have successfully carried out the experiment using the proton and carbon spins of trichloroethylene, and confirmed the result using state tomography. We have thus extended the space of entangled quantum states explored systematically to three quantum bits, an essential step for quantum computation.

We live in a world which, to the best of our knowledge, is surprisingly well described by the laws of quantum mechanics. Fundamental processes in nature are all compatible with quantum mechanics, even if some of its predictions are counterintuitive.

A good example of this peculiar behavior is given by the now famous pairs of two state systems described by Einstein, Podolsky and Rosen (EPR) [1]. The unorthodox behavior of this composite system has been crystallized by the Bell inequalities [2], which give a statistical test for the existence of *elements of reality* [3]. In our world, photons in a polarization state of the form $(|00\rangle - |11\rangle)/\sqrt{2}$ violate these inequalities as shown by Aspect et al. [4] and therefore contradict the existence of elements of reality.

In a beautiful paper, Greenberger, Horne and Zeilinger (GHZ) [5] demonstrated that it is possible, in a single run of experiments using a state of the form $|000\rangle + |111\rangle$ (the GHZ state), to refute the existence of elements of reality. The GHZ experiment has been succinctly summarized by Mermin [6]. Even though this experiment is very appealing, nobody has yet been able to perform it. The reason is that it is rather difficult to precisely manipulate entangled quantum states of many particles. In fact, up to now, entangled pure states of only two particles have been systematically explored [4, 7, 8, 9].

In this letter we show how three particle entangled states can be realized using nuclear magnetic resonance techniques. We have carried out the experiment and verified that indeed we had a GHZ state by using tomography [10]. We first describe how it is possible to obtain a pure state result from the initial mixed density matrix of a NMR system. Then we explain the sequence of operations used to obtain a GHZ state and finally we give the experimental results.

We will not investigate the non-local behavior of GHZ states, as NMR is not appropriate for this undertaking. What we have created, however, is a typical

state needed for a three-bit quantum computer. Indeed our approach is the one used for quantum computation and we refer the reader to [11] for an introduction.

The usual approach for quantum computation requires an initial pure state [11]. The computation itself is a transformation of this initial state using unitary operations. Useful information is then extracted by a measurement of the final state. However, it has recently been shown that the computation can be performed using an initial mixed state as long as the decoherence time is sufficiently long [12,13,14,15].

In liquid state NMR, the computation takes place on a large ensemble of identical quantum systems. Each member of the ensemble of quantum systems consists of the interacting nuclear spins of a molecule in a high magnetic field. The initial state of the nuclear spins is achieved by allowing the system to relax to thermal equilibrium. Information processing with such an ensemble can be divided into three steps consisting of preparation, computation and readout. Each of these steps is equivalent to an application of certain quantum operations identically to each member of the ensemble.

The nuclear spins are manipulated by applying radio frequency (RF) pulses tuned to the Larmor frequencies of the spins [16]. The spins can be selectively excited by exploiting differences in Larmor frequencies. Entanglement and two spin operations are achieved by a delay to allow for interaction between spins. In our case, these interactions are scalar couplings which can be selectively turned off by the use of refocusing pulses tuned to one of the nuclei. Any unitary quantum operation can be decomposed into such operations [17,18,19].

The measurement step in NMR consists of observing the signal induced in RF coils by the precession of the nuclear spins [16]. In effect, we measure the time evolution of the expectation of the σ_x and σ_y operators for each nuclear spin. By Fourier transforming the signal and analyzing the spectrum, other operators such as tensor products of either σ_x or σ_y with I or σ_z for pairs of interacting spins can be measured. The traceless part of the density matrix (called the deviation matrix) can be determined by a tomography procedure [10]. This involves several measurements of spectra after applying different reading pulses to the final state, thus permitting observation of all elements of the deviation matrix.

The main problem in using NMR for the observation of multiparticle entanglements and quantum computation is that the sample is initially in a highly mixed state. There are several methods for overcoming this limitation without cooling the sample. One idea is to transform the initial mixed state so that we have a *pseudo-pure* state

$$\rho_{pp} = p |0 \dots 0\rangle\langle 0 \dots 0| + \frac{1-p}{N} I, \quad (1)$$

where N is 2^n for n qubits and $p + (1-p)/N$ is the probability of the ground state ($|0 \dots 0\rangle$). In NMR this state is indistinguishable from a pure state because all observables are traceless. The methods discussed in the literature [12,13,14,15] are all based on the idea of viewing the non-ground state components of the initial thermal density matrix as noise and applying an averaging method to eliminate their contribution to measurements. These methods are in principle

sufficient for exploiting the advantages of quantum computing in an ensemble setting. In our experiments we used a new method for extracting pseudo-pure states from experimental data. The method is based on two features of our experiment. The first is that the sequence of pulses and delays applied to the sample for preparing the GHZ state are much shorter than the decoherence time for the nuclear spins (about 5ms versus at least 210ms). As a result, the evolution is very close to unitary and therefore preserves the eigenvalue structure of the initial state. This structure is well known and determined by the thermal distribution. The unitarity property can be tested by comparing the eigenvalues of the output state to those of the thermal distribution. The second feature is that complete state tomography is used to analyze the output of the experiment. This allows us to decompose the output state into its eigenstates. The state $|000\rangle$ is associated with the smallest eigenvalue of the thermal state and must be transformed by unitary evolution into the corresponding eigenstate of the output. By comparing this eigenstate to the desired GHZ state, we learn whether our pulses indeed generated this state from that component of the ensemble which was initially in $|000\rangle$.

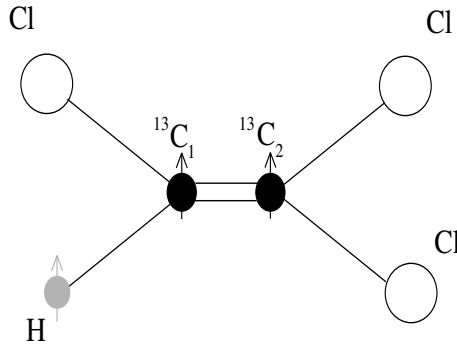


Fig. 1. Pictorial representation of trichloroethylene. The three qubits are given by the nuclei of hydrogen (H) and of the two carbon ^{13}C 's (C_1 and C_2). The last two are distinguishable because of the asymmetry of the chlorine environment. The chlorine nuclei have spin $3/2$ and their interaction with the qubits can be neglected.

In order to create a GHZ state we need three nuclei. A convenient system is trichloroethylene, the molecule shown in Figure 1. The spins of the hydrogen and carbon nuclei were used for the three quantum bits. They interact only weakly with the chlorine nuclei, which can therefore be ignored. In a strong static magnetic field (11.5 Tesla along the z -axis) the evolution of the hydrogen and the two carbon nuclei are well described by the Hamiltonian

$$H = -\omega_H \sigma_z^H - \omega_{C_1} \sigma_z^{C_1} - \omega_{C_2} \sigma_z^{C_2} + J_{HC_1} \sigma_z^H \sigma_z^{C_1} + J_{C_1 C_2} (\sigma_x^{C_1} \sigma_x^{C_2} + \sigma_y^{C_1} \sigma_y^{C_2} + \sigma_z^{C_1} \sigma_z^{C_2}) + J_{HC_2} \sigma_z^H \sigma_z^{C_2} \quad (2)$$

with $\omega_H \approx 500.1334915$ MHz, $\omega_{C_1} \approx 125.7725805$ MHz, $\omega_{C_2} \approx 125.7732305$ MHz, which gives a chemical shift of 650 Hz between the carbons. The J couplings have values of $J_{HC_1} \approx 203$ Hz, $J_{C_1C_2} \approx 102$ Hz and $J_{HC_2} \approx 10$ Hz.

The data was acquired with a Bruker DRX-500 spectrometer using doubly labeled trichloroethylene ($^{13}C_1, ^{13}C_2, 99\%$). The relaxation time (T_1) for the hydrogen is 7s and 30s for the carbons. The phase decoherence time (T_2) for hydrogen is 3s and 0.4s and 0.2s for C_1 and C_2 respectively.

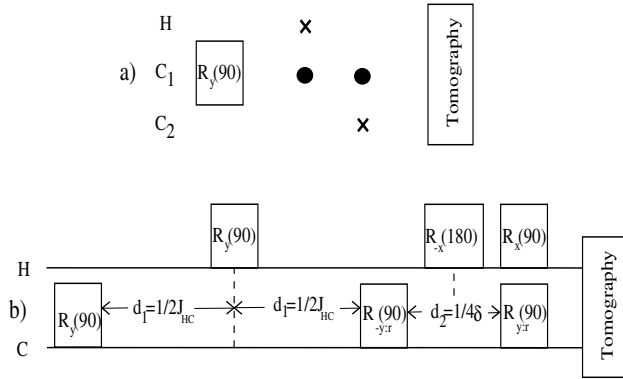


Fig. 2. a) Circuit to create a GHZ state. $R_a(\theta)$ corresponds to a rotation of the qubit by an angle θ around the a -axis and the two other gates are CONTROL-NOTs (see DiVincenzo[14]). b) Implementation by an NMR pulse sequence producing a GHZ state (in the rotating frame of H and C_1). The delay d_1 serves to let the coupling between the nuclei create the entanglement, and d_2 is used to generate a phase shift on the C_2 nucleus. Coupling between the carbon nuclei during this delay is negligible. All pulses are non-selective.

A simple circuit to create a GHZ state consists of a rotation by $\pi/2$ around the y -axis followed by two CONTROL-NOT gates on the other qubits. Adapting this to our system results in the pulse sequence shown in figure 2 b). The deviation density matrix, initially given by the state

$$\rho_{\Delta}^i = \omega_H \sigma_z^h \otimes 1 \otimes 1 + \omega_{C_1} 1 \otimes \sigma_z^{C_1} \otimes 1 + \omega_{C_2} 1 \otimes 1 \otimes \sigma_z^{C_2} \quad (3)$$

is transformed to

$$\begin{aligned} \rho_{\Delta}^f = & -\omega_H \sigma_z^h \otimes \sigma_z^{C_1} \otimes 1 - \omega_{C_1} \sigma_x^h \otimes \sigma_x^{C_1} \otimes 1 \\ & - \omega_{C_2} 1 \otimes \sigma_z^{C_1} \otimes \sigma_z^{C_2} \end{aligned} \quad (4)$$

The tomography procedure was implemented with twelve sets of reading pulses. For each set, two experiments were performed to read off the hydrogen and the carbon spectra.

To compute the peak intensities for each spectrum, we assumed that each peak is approximately Lorentzian. The peak positions and decay parameters associated with the Lorentzian shape were obtained by optimizing (in the least squares sense) the match to the calibration spectra. The carbon and hydrogen spectra were both matched to within 2%. This should be compared to the estimated noise, which was determined to be less than .5% for the carbon and .01% for the hydrogen nuclei. The mismatch is due to additional peaks, primarily from unlabeled compound, and also to shimming problems causing deviation from the ideal Lorentzian lineshape. The decay parameters (essentially T_2^*) obtained after optimization were $0.51 \pm 0.03\text{sec}$ for the hydrogen nucleus, $0.41 \pm 0.01\text{sec}$ for carbon 1 and $0.23 \pm 0.01\text{sec}$ for carbon 2 (the error bars are estimated from the variation between different peaks of the same nucleus). The latter agree well with the experimentally determined T_2 time for the carbons. Since T_2 for the hydrogen is near 3sec, it can be seen that substantial peak broadening due to magnetic field inhomogeneity reduces the hydrogen T_2^* .

After the peak positions and decay parameters were determined, each spectrum of the experiment was deconvolved as a linear combination of the ideal peaks. The coefficients of the combination yield the intensity and phase of each peak in the spectrum. These numbers were then normalized by the calibration intensities and phase corrected using the calibration phases. Then they were used to determine the deviation density matrix of the output state of the experiment.

An idea of how *unitary* the evolution has been is obtained by comparing the eigenvalues of the initial deviation matrix ($\{24, 16, 16, 8, -8, -16, -16, -24\}$) to the final one ($\{24.1, 16.3, 15.5, 7.7, -8.1, -15.0, -16.5, -23.8\}$). Clearly the transformation was unitary to a very good approximation.

Using the procedure explained above we get, after diagonalizing and taking the appropriate eigenvector, the experimentally determined GHZ state

$$\rho_e^{GHZ} = 10^{-3} \begin{bmatrix} 493 & -53 & -47 - 18i & -28 - 5i \\ -53 & 6 & 5 + 2i & 3 \\ -47 + 18i & 5 - 2i & 5 & 3 \\ -28 + 5i & 3 & 3 & 2 \\ 25 + 22i & -3 - 2i & -2 - 3i & -1 - 2i \\ -14 - 31i & 2 + 3i & 3i & \\ -31 + 13i & 3 - i & 3 & 2 \\ 468 + 147i & -51 - 16i & -39 - 31i & -25 - 13i \\ 25 - 22i & -14 + 31i & -31 - 13i & 468 - 147i \\ -3 + 2i & 2 - 3i & 3 + i & -51 + 16i \\ -2 + 3i & -3i & 3 & -39 + 31i \\ -1 + 2i & -2i & 2 & -25 + 13i \\ 2 & -2 & -2i & 30 + 14i \\ -2 & 2 & 2i & -23 - 25i \\ 2i & -2i & 2 & -25 + 21i \\ 30 - 14i & -23 + 25i & -25 - 21i & 488 \end{bmatrix}. \quad (5)$$

A pictorial representation is given in figure 3. The fidelity of the state compared to the ideal GHZ state is

$$\mathcal{F} = \langle \Psi_{GHZ} | \rho_e^{GHZ} | \Psi_{GHZ} \rangle = 0.95. \quad (6)$$

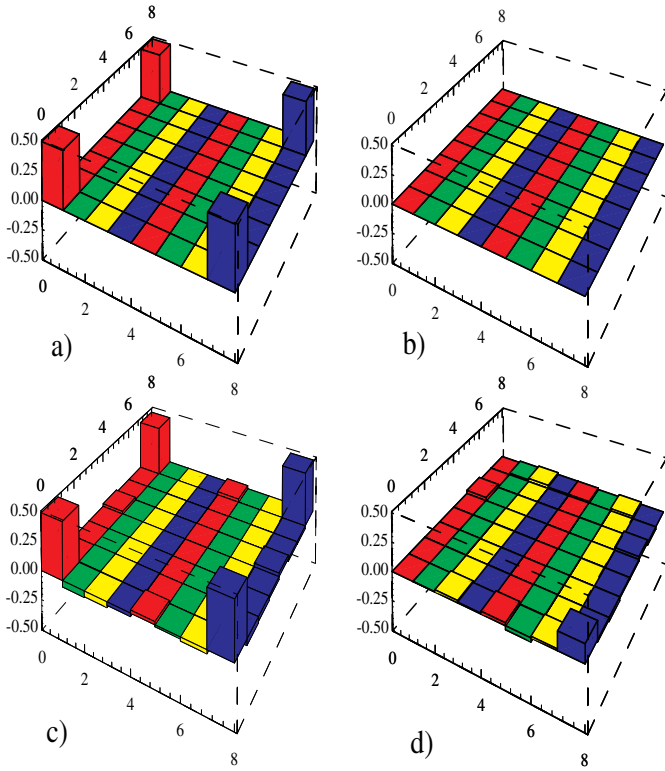


Fig. 3. Pictorial representations of the theoretical and experimental density matrices ρ for the GHZ state $|000\rangle + |111\rangle$. a), b) Real and imaginary parts of the theoretically determined density matrix. c), d) Real and imaginary parts of the experimentally measured density matrix. Each bar graph represents the value of the transitions between the eight states $|000\rangle \dots |111\rangle$ by the height of the bar in the corresponding position of the 8×8 array.

In conclusion, we have shown how to construct an effective GHZ state with a fidelity of 95% in NMR starting with the thermal mixed state. The experiments demonstrate the ability in NMR to fully explore the state space of multi-particle systems, which is all that is required for quantum computation. By using a four spin system, the paradoxical output of the originally proposed GHZ experiments can be observed as proposed in [20]. However, due to the microscopic separation of the particles involved and the method for observation used, this would not be a true test of the existence of elements of reality.

To construct our GHZ state we developed a new method for extracting pseudo-pure states from NMR spectra. The method can be used efficiently for process tomography [10], since each experiment determines the transformation

of the state space induced by the applied operation on each of the eigenstates of the input state.

Our experiments demonstrate that room temperature liquid NMR is well suited for quantum computations involving small numbers of qubits. Although manipulating three qubits is a small step for large scale quantum computation, it is the first time that a quantum network has been used to systematically entangle more than two qubits.

Acknowledgments. We would like to thank David Cory and Timothy Havel for useful conversations and the Stable Isotope Laboratory at Los Alamos for the use of their facility. This research was supported in part by the National Security Agency.

References

1. A. Einstein, B. Podolsky, N. Rosen, *Phys.Rev.* **47**, 777 (1935).
2. J.S. Bell, *Physics* 1, 195 (1964).
3. EPR define elements of reality as: "If, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity" in [1].
4. A.Aspect, P. Grangier and G. Roger, *Phys.Rev.Lett.* 47,460, 1981
5. D.M. Greenberger, M. Horne, A. Zeilinger, in *Bell's Theorem, quantum Mechanics, and Conceptions of the Universe*, M. Kafatos, ed., Kluwer, Dordrecht, The Netherlands (1989), .p.69.
6. D. Mermin, *Physics Today*, June 1990, p9.
7. C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano and D.J. Wineland, *Phys.Rev.Lett.* 75, 4714, 1995.
8. Q. A. Turchette and C. J. Hood and W. Lange and H. Mabuchi and H. J. Kimble, *Phys.Rev.Lett.* 75, 4710, 1995.
9. H. Walther, Single atom experiments in cavities and trap, submitted to *Proc. Roy. Soc.*, 1997.
10. M. Raymer, M. Beck abd D. McAllister, *Phys.Rev.Lett.* 72,1137, 1994.
11. D. DiVincenzo, *Science*, 270, 255, 1995.
12. D. G. Cory, A. F. Fahmy, and T. F. Havel, *Proceedings of the National Academy of Sciences of the United States of America* **94**, 1634 (1997).
13. D. G. Cory, M. Price, A. F. Fahmy, and T. F. Havel, Nuclear magnetic resonance spectroscopy: an experimentally accessible paradigm for quantum computing, *Physica D* in press.
14. N. A. Gershenfeld and I. L. Chuang, *Science* **275**, 350 (1997).
15. E. Knill, I. Chuang & R. Laflamme, Effective pure states for bulk quantum computation, quant-ph/970653.
16. R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford University Press, Oxford, 1994).
17. A. Barenco *et al.*, *Phys. Rev. A* **52**, 3457 (1995).
18. S. Lloyd, *Phys. Rev. Lett.* **75**, 346, 1995 .
19. D. P. DiVincenzo, *Phys. Rev. A* **51**, 1015, 1995.
20. S. Lloyd, A Greenberger-Horne-Zeilinger experiment for mixed states, quant-ph/9704013.

Quantum Computing Using Electron-Nuclear Double Resonances

Charles M. Bowden⁽¹⁾, Jonathan P. Dowling⁽¹⁾ and Steven P. Hotaling⁽²⁾

- (1) Weapons Sciences Directorate, AMSMI-RD-WS-ST, Missile Research, Development, and Engineering Center, U. S. Army Missile Command, Redstone Arsenal, Alabama 35898-5248
- (2) U. S. Air Force Rome Laboratory, Rome Laboratory, 26 Electronic Parkway, Rome, New York 13441-4514

Abstract. We consider the use of Electron-Nuclear Double Resonance (ENDOR) techniques in quantum computing. ENDOR resolution as a possible limiting factor is discussed. It is found that ENDOR and double-ENDOR techniques have sufficient resolution for quantum computing applications.

1. Introduction

Recently, one of the authors presented arguments that the Electron Nuclear Double Resonance (ENDOR) process may be exploited for the storage and processing of discrete data on a quantum scale¹⁻⁶. The implication is that a solid state realization of a quantum mechanical computer could be engineered. Among the advantages are that such a solid state quantum computer would be stable, programmable, and input/output (I/O) controllable by current state-of-the-art technology. It could be envisioned, in principle, to be “engineerable” for considerably lower per unit cost than quantum computers operating on the principle of induced quantum superposition and entangled states of trapped ions⁷, or photon states using microcavities⁸. Furthermore, solid state ENDOR is a well established procedure (established by Feher⁹ in 1959), and laboratory components are commercially available at reasonable costs. The present paper is a proposed novel quantum computing paradigm based upon the use of multipulse resonance techniques to manipulate nuclear spins of a mostly relatively low dimensional ensemble deviation from thermal equilibrium¹⁰. The new paradigm builds upon a previously proposed paradigm which utilizes well established techniques from nuclear magnetic resonance (NMR) spectroscopy¹¹. An obvious advantage to using superposition of nuclear spins for quantum logic gates, and nuclear spin flips to conduct quantum computing, is the possibility of extraordinary long decoherence times due to the relative isolation of nuclear spins within a molecule. A disadvantage is related to this, and that is the length of time required to couple information in and out of a system and manipulation during computation. Another disadvantage is that nuclear spin flips are induced at radio frequency (rf) wavelengths, and so quantum computation is restricted to temporal unitary evolution and is entirely non-local.

In the present paper, we propose a scheme based upon Electron Nuclear Double Resonance (ENDOR) as a means to practical quantum computation. The proposed scheme builds from the previous proposals which use NMR spectrographic techniques^{10,11}, but trades reduction in decoherence time by electron spin, nuclear spin coupling, but gains in high I/O bit rates and stronger coupling to manipulate computation. Also, we shall point out that sequential spatially dependent architecture is possible using laser electronic excitation to manipulate electron spin coupling to nuclear spins.

A brief discussion of Electron Spin Resonance (ESR) and ENDOR will be presented in the next section. Our novel paradigm for quantum computing using laser-induced electronic excitation and ENDOR will be discussed in Section 3, and Section 4 will be used for summary and conclusion.

2. ESR and ENDOR Background

The theory of ESR is derivable from the Dirac theory and will not be treated here, but may be found in the literature^{1-6, 9, 12-13}. Here, we simply state the essential spin Hamiltonian and discuss its interaction terms,

$$Ham = \beta \mathbf{H} \cdot \mathbf{g} \cdot \mathbf{S} - \beta_N \mathbf{H} \cdot \mathbf{g}_N \cdot \mathbf{I} + \lambda \mathbf{L} \cdot \mathbf{S} + \mathbf{S} \cdot \mathbf{A} \cdot \mathbf{I} \quad (1)$$

The first and second terms correspond to the Zeeman energy contributions due to the electron spin, \mathbf{S} , and nuclear spin, \mathbf{I} , coupling to the magnetic field \mathbf{H} . Here, \mathbf{g} is the coupling tensor in units of the Bohr magneton β and \mathbf{g}_N is the nuclear spin coupling tensor in units of the nuclear magneton, β_N . The third term in the Hamiltonian is the Zeeman interaction associated with the coupling between the magnetic moment due to the electron's intrinsic spin and that due to the electron's orbital momentum in a bound state. This term contains information about electronic defect states with different microscopic properties through measurement of shifts in g-value of a paramagnetic site. The last term is the hyperfine term which expresses the magnetic interaction between the nuclear and electronic spins due to overlap of the electronic wave functions with nuclear spins. This term, governed by the interaction tensor \mathbf{A} , depends upon nuclear spin contact interaction with electronic wave functions and can be nearly isotropic, as with nearly s-type electronic orbitals, or anisotropic as for p- or d-like orbitals. By analyzing the energy contributions in the hyperfine term, the nature of the spin center (type of electronic state) can be determined.

Electron Nuclear Double Resonance (ENDOR) provides the capability to more closely examine the anisotropic hyperfine interaction in terms of the atomic and electronic interactions at the paramagnetic centers. Ions or free radicals trapped in a solid lattice experience perturbations in their energy levels as expressed by this matrix. These perturbations can affect the spin transition dynamics of the paramagnetic species, and be detectable by Electron Spin Resonance (ESR). ENDOR allows the hyperfine and spin lattice relaxation phenomena to be measured by detecting the Nuclear Magnetic Resonance (NMR) signal as a change in the ESR spectrum.

In ENDOR, the nuclear spins are modulated by addition of a transverse rf field, while the electron spins are driven by a transverse microwave (MW) field. A

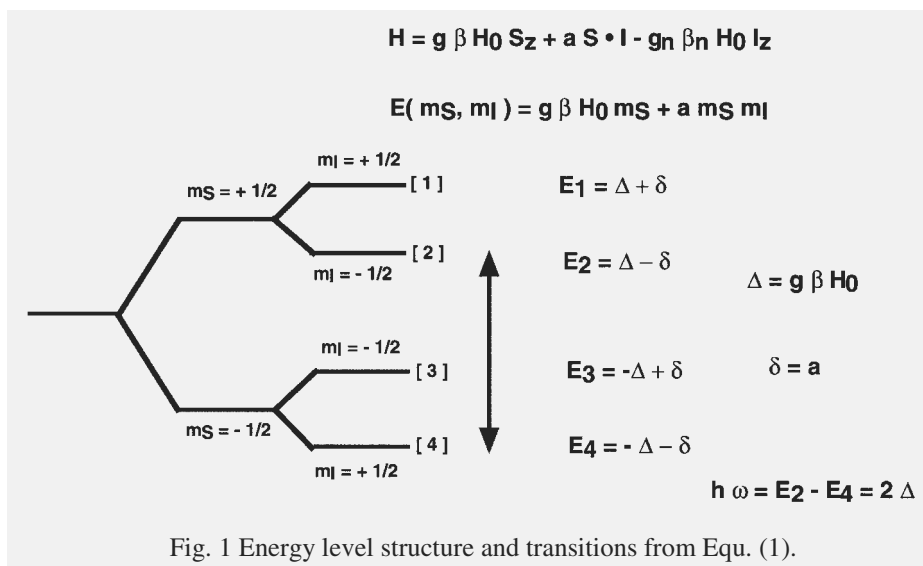
simplified version of the ENDOR process can be illustrated by the simplified version of Eq. (1),

$$H_{\text{am}} = g \beta H_0 S_z + a \mathbf{S} \cdot \mathbf{I} - g_N \beta_N H_0 I_z \quad (2)$$

where we have neglected spin-orbit coupling, which in many cases is quenched¹². The eigenenergies associated with H, Eq. (2), in terms of the appropriate quantum numbers are given by

$$E(m_S, m_I) = g \beta H_0 m_S + a m_S m_I \quad (3)$$

where we assume that the electronic Zeeman energy, given by the first term, and the hyperfine energy, the second term in Eq. (3), are much larger than the nuclear Zeeman energy. The corresponding energy level structure, together with the ordering of the energy levels and transitions in terms of the quantum numbers is illustrated in Fig. 1.



Provided $\delta \neq 0$, transitions $|4\rangle \rightarrow |2\rangle$ and $|3\rangle \rightarrow |1\rangle$ cause simultaneous electron-nuclear double spin flips at the transition energy $-\omega = 2\Delta$. Whereas, transitions $|4\rangle \rightarrow |1\rangle$ and $|3\rangle \rightarrow |2\rangle$ correspond to electron spin flips only, but at the transition energy $-\omega = 2\Delta + 2\delta$. The usual ENDOR procedure requires that the electronic transition be saturated using a microwave field at frequency $\omega = 2\Delta / \hbar$; then nuclear spin flips are induced with an rf field at frequency $\Omega = 2\delta / \hbar$ and appear as modifications in the electron spin-resonance spectrum.

To date, the most prevalent application of the ESR process in solid state materials is the determination and characterization of defect structures. As discussed above, the hyperfine interaction between the magnetic moments of an unpaired electron and neighboring nuclei can yield this information. The hyperfine interaction is sometimes not well resolved due to lattice phonon modes. This is especially true of the super-

hyperfine interaction or ligand hyperfine interaction that appears in the ESR spectrum as an interaction between the magnetic moment of an unpaired spin and its nearest neighbor nuclei. To the untrained eye, this broadening effect appears as a fundamental resolution limitation of magnetic resonance techniques. However, in the ENDOR process, the NMR transitions of neighboring nuclei interacting with the unpaired spin are measured by detecting their influence on the unpaired spin's polarization under favorable signal-to-noise experimental conditions (partially saturated spin-resonance condition). These ENDOR-detected NMR transitions are detected as quantum mechanical transitions of much higher energy than would be observed in conventional NMR or ESR techniques. This implies that there are far fewer lines in the ENDOR spectrum to resolve than in the conventional ESR or NMR spectra. This is shown schematically in Figure 2 for the spin 1/2 system where we note that the ENDOR transition (cross transition) is larger than, and fewer in number than, either the ESR or NMR transitions. The effect is more noticeable for higher order spin systems ($I = 3/2, \dots$, etc.). This enhancement has been reported experimentally in SrFCl and BaFCl systems^{14–16}.

A strong advantage associated with the paradigm presented here is the possibility to control the transferred hyperfine interaction, the tensor **A** in Eq. (1), used as a generic example and illustrated in Fig. 3a. Here, an unpaired electron orbital associated with atom, *A*, is represented as an s-state in the electronic ground state, without any overlap of the wave function at the nucleus of atom, *B*. Laser excitation of atom, *A*, on the other hand, induces an electronic transition to a p-orbital or d-orbital, with consequent overlap of the electron wave function at the nucleus, *B*, inducing S_A, I_B electron spin, nuclear spin interaction, as shown in Fig. 3b. If a microwave field is tuned to the transition shown in Fig. 1, a simultaneous electron spin S_A , nuclear spin, I_B , transition, double spin flip is induced. Thus, nuclear spin flips can be controlled in atom, *B*, by electron spin flips controlled by laser-induced transferred hyperfine interaction¹⁷.

This scheme constitutes a significant modification of the NMR quantum computation of Refs. 10–11. Here, we use that paradigm to build a controlled NOT-gate conditional on a reference nuclear spin, but introduce laser field-induced electron spin, nuclear spin coupling by laser-induced transferred Fermi contact interaction. Thus, input/output and control can be executed locally under unitary time evolution, $U = \exp(-i \text{Ham} \Delta t / \hbar)$, and the laser field can be used to induce π -pulse excitation/de-excitation of duration Δt in sub-picosecond time scale. Thus, decoherence can be minimized and control can be executed on the ultrafast time scale. In essence, we trade off some decoherence for fast local control.

It is also possible to utilize Double-ENDOR (D-ENDOR) effects to increase the spectral resolution. In D-ENDOR, two NMR transitions are stimulated while the ESR transition is measured. In this case, typical improvements of a factor of 10:1 in resolution are obtained¹⁸. In addition, Optical Detected Magnetic Resonance (ODMR) has the potential for increased output resolution.

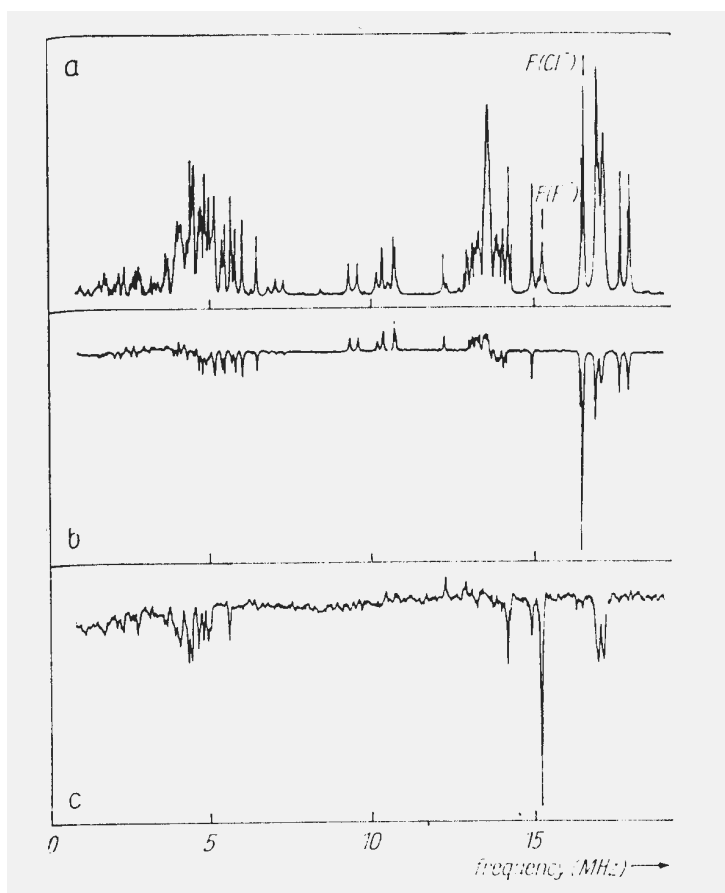


Fig. 2 a) EPR spectra of $F(Cl^-)$ and $F(F^-)$ centers in $BaFCl$ for $B0$ 150 off c in the c - a plane. b) Double ENDOR spectrum for setting $f1$ to an ENDOR line of $F(Cl^-)$ centers (see mark on Fig. 2a). c) Double Endor spectrum for setting $f1$ to an ENDOR line of $F(F^-)$ centers (see mark in Fig.2a). [From Reference 15, J. R. Niklas, R. U. Bauer, and J. M. Spaeth, *Phys. Stat. Sol. (b)* **119**, 171 (1983).]

3. Quantum Logic Circuits

Recently, quantum mechanical Hamiltonians for logic gate elements have been derived for: NOT¹⁸, exclusive OR^{19,20}, and controlled-NOT (C-NOT)¹⁰ operations. The former papers concerned theoretical derivations while in the latter paper by Gershenfeld and Chuang, NMR transitions were experimentally demonstrated to realize a C-NOT operation in hardware. This experiment, when taken along with earlier work^{1-6, 18-20}, leads us to propose an ENDOR-based process paradigm for realization of higher complexity quantum logic gates.

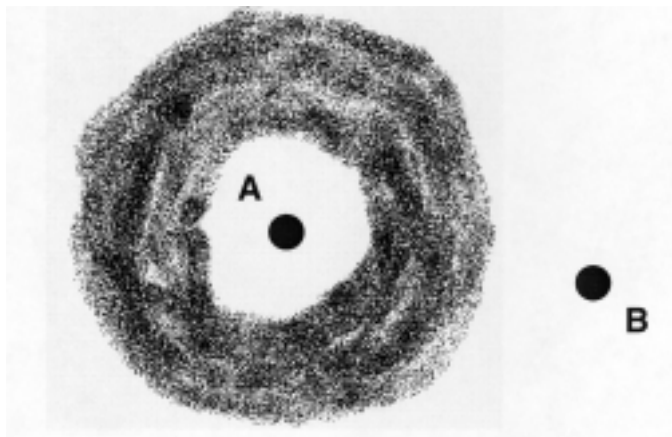


Fig. 3a Electronic S - state; $A=0$, ground state, no overlap.

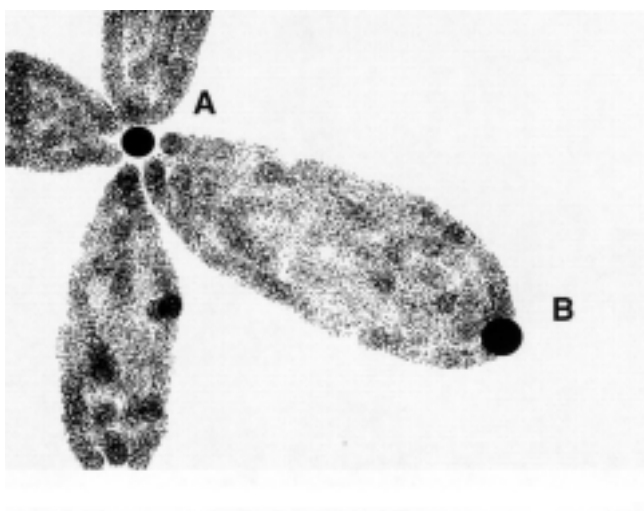


Fig. 3b Electronic p - state; A nonzero, excited state, overlap.

Consider a spin system consisting of a “free” spin (electron or hole) in some photoactive crystal or polymer, its nearest neighbor nucleus B and its second nearest neighbor nucleus C , as illustrated in Figure 4. The magnetic moments of B and C are assumed distinct. Spin-spin interaction between S_A and I_B and S_A and I_C is defined

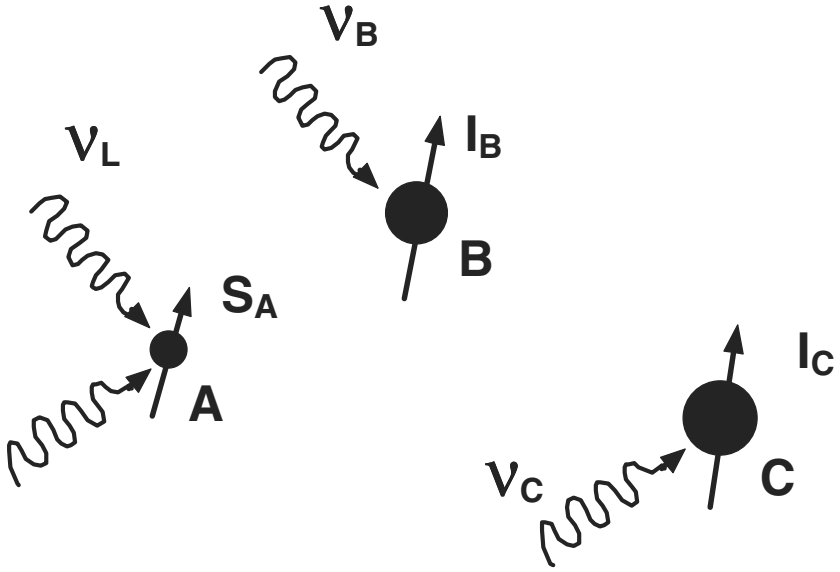


Fig. 4 Double ENDOR system; see text.

by the anisotropic hyperfine tensor A [Eq. (1)]. Stimulating nuclei B and C by separate rf wavelengths v_B and v_C corresponding to NMR transition frequencies while simultaneously stimulating S_A by microwave radiation sets up a D-ENDOR system. Stimulation of S_A by external laser radiation, v_L , serves as a control to flip the spin states of the D-ENDOR system. Alternately, I_B or I_C could be perturbed by IR lasers, but for simplicity, we assume only S_A laser stimulation. This controlled stimulation thus realizes a controlled NOT gate. Consider the spin interaction to be coupled to a chain of atoms (D, E, \dots, N). Repeated pulsing of S_A would cause the quantum chain to respond as the quantum mechanical analog of a chain of emitters as shown in Figure 5. If N is even, then there is no net change in spin state at output, or $|\text{up}\rangle \rightarrow |\text{up}\rangle$. If N is odd, $|\text{up}\rangle \rightarrow |\text{down}\rangle$.



Fig. 5 Spin coupled to a chain of atoms; N even, no change at output, N odd, spin flip at output. Constitutes a controlled NOT gate.

4. Conclusion

We have proposed a new paradigm for quantum computing which begins with the construction of a quantum controlled- NOT gate as prescribed in Reference 10, which uses a non-equilibrium ensemble of nuclear spins. The experiments of Gershenfeld, *et al.* have demonstrated exceptionally long decoherence times due to the relative isolation of nuclear magnetic moments to externally induced transitions. Our scheme introduces a higher-order process by coupling electron, nuclear double resonant super-hyperfine transitions, controlled by laser-induced electronic transitions. The advantages of this scheme are that each complex molecule becomes, in and of itself, a quantum computer and the entire system represents massive parallelism. In addition, we have shown that laser-induced electronic excitation renders local control of gate preparation, spin flips, and input/output which can take place on the sub-picosecond time scale. Here, we trade long decoherence times, intrinsic with nuclear spin flips, with controlled coupling with electron spin flips by transferred or super-hyperfine interactions. The latter is regulated, i.e., on or off, by laser field π -pulse electron excitation/de-excitation. We hope that this will lead to near-term experimental investigations.

5. References

1. S. P. Hotaling, "The Influence of Transition Metal Dopants on the Properties of Bismuth Metal Oxide (BM)) Sillenites Grown by the Czochralski and Hydrothermal Techniques," Ph.D. thesis, Clarkson University, 1995.
2. S. P. Hotaling, "Photonic Excitations in Bismuth Silicon Oxide," in *Photonic Device Engineering for Dual Use Applications*, Andrew R. Pirich, ed., Proc. SPIE **2481**, 240–247, 1995.
3. S. P. Hotaling, "Photon-Spin Interactions: A Potential Foundation for Photonic Quantum Computing, Proc. SPIE **2487**, 1996.
4. S. P. Hotaling, "Radix $R > 2$ Quantum Computation," Proc. International Quantum Structures Conference, Berlin, Germany, 1996.
5. S. P. Hotaling and A. R. Pirich, "General Purpose Quantum Computation," U. S. patent application, submitted, 1997.
6. S. P. Hotaling and A. R. Pirich, "Radix $R > 2$ Quantum Computation," U. S. patent application, submitted, 1997.
7. C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, "Demonstration of a Fundamental Quantum Logic Gate," Phys. Rev. Lett. **75**, 4714–4717, 1995.
8. Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuci, and H. J. Kimble, Phys. Rev. Lett. **75**, 4710–4713, 1995.
9. G. Feher, Phys. Rev. **114**, 1219–1249, 1959.
10. N. A. Gershenfeld and I. L. Chuang, "Bulk Spin-Resonance Quantum Computation," Science **275**, 350–356, 1997.
11. D. S. Cory, A. F. Fahmy, and T. F. Havel, "Ensemble Quantum Computing by NMR Spectroscopy," Proc. National Academy of Sciences, USA **94**, 1634–1639, 1997.

12. C. P. Slichter, *Principles of Magnetic Resonance*, (Harper & Row, 1963).
13. C. P. Poole and H. A. Farach, *Handbook of Electron Spin Resonance* (AIP Press, 1994).
14. R. U. Bauer, J. R. Niklas, and J. M. Spaeth, *Phys. Status Solidi B* **118**, 557, 1983.
15. J. R. Niklas, R. U. Bauer, and J. M. Spaeth, *Phys. Status Solidi B* **119**, 171, 1983.
16. R. U. Bauer, J. R. Niklas, and J. M. Spaeth, *Radiat. Effects* **287**, 1983.
17. C. M. Bowden and J. E. Miller, "Superhyperfine Interaction in the Electron-Spin Resonance Spectrum of Substitutional Gd^{3+} Impurity in CaF_2 Single Crystals Under Applied Stress," *Phys. Rev. Lett.* **19**, 4, 1967.
18. D. Mozyrsky, V. Privman, and S. P. Hotaling, "Design of Gates for Quantum Computation: The NOT Gate," *Quant-ph*, 9608029, Los Alamos National Laboratory, E-Print, 1996.
19. D. Mozyrsky, V. Privman, and S. P. Hotaling, "Design of Gates for Quantum Computation: The Three Spin XOR in Terms of Two-Spin Interactions, Los Alamos National Laboratory, E-Print, *Quant-ph* 9612029, 1996.
20. D. Mozyrsky, V. Privman, and S. P. Hotaling, "Extended Quantum XOR Gate in Terms of Two-Spin Interactions," Los Alamos National Laboratory, E-Print, *Quant-ph* 9610008, 1996.

Physical Implementations for Quantum Communication in Quantum Networks

H.-J. Briegel^{1,2}, J.I. Cirac^{1,2}, W. Dür¹, S.J. van Enk³, H.J. Kimble³,
H. Mabuchi³, and P. Zoller¹

¹Institut für Theoretische Physik, University of Innsbruck, A-6020 Innsbruck, Austria

²Universidad de Catilla-La Mancha, 13071 Ciudad Real, Spain

³Norman Bridge Laboratory of Physics, California Institute of Technology 12-33,
Pasadena, CA 91125

Abstract. An overview of our work on quantum communication in quantum networks is given. We discuss physical implementations for quantum networks based on cavity-QED, including error correction schemes for the transmission of qubits over noisy channels and for noisy local entanglement and joint measurement operations on atoms. We define the "photonic channel", and outline the most recent developments concerning the "quantum repeater".

1 Introduction

Quantum Communication requires the ability to prepare, manipulate and measure single quantum bits. Here we will discuss a quantum-optical implementation for Quantum Communication based on elements that have been realized experimentally, or that can be expected to be realized in the near future: trapped ions or atoms, and cavity quantum electrodynamics (CQED) both in the optical and microwave regime [1]. For instance, among the recent experimental highlights in these fields are the first steps towards the implementation of *quantum-logical operations*, such as two-qubit operations that form the building blocks for quantum computing. Thus, one has performed a universal 2-bit quantum gate, using the quantized center-of-mass motion of a trapped ion and its internal quantum degrees of freedom [2]. In CQED the basic elements of gate operations involving the internal state of an atom and a cavity mode have been demonstrated [3], and recently one has succeeded in entangling internal states of two atoms [4].

Given this progress we expect to see in the lab within the next few years small or mesoscopic quantum (optical) systems consisting of a few particles (quantum bits). It is, therefore, important to identify applications of quantum mechanics and quantum information which can be realized with finite resources in direct relation to these new coming generations of experiments. In the present paper we will discuss the possibility of implementing quantum communication in (small) quantum networks. Such a quantum network involves nodes in which qubits are stored and locally processed. These nodes are connected by quantum wires (such as an optical fiber) over which communication takes place by sending photons.

Such networks are also interesting from the point of view of connecting quantum computers and implementing a model of distributed quantum computing. For storing and manipulating quantum information atoms are perfect candidates. The qubit is stored in two ground states (one or both states could be metastable states as well) $|0\rangle$ and $|1\rangle$, so that the quantum memory does not suffer from spontaneous emissions. In order to place the atom in a well-defined environment, a cavity is used and the atom is strongly coupled to one cavity mode.

The basic idea for transmitting a qubit is now to manipulate the atom by appropriately designed laser fields such that the qubit stored in the atom is coherently transferred to the cavity mode. Then the cavity photon will leak out of the cavity, travel through vacuum or a fiber to another cavity, enter that cavity, and by laser manipulation the information can finally be transferred to another atom. In [9] the problem of constructing laser pulses that fulfill this goal has been solved. In [15] a different idea was proposed, based on adiabatic passage, so that the requirement on pulse shapes is relaxed. Crucial here is also the possibility of implementing error correction schemes with finite resources. Standard error correction schemes and purification protocols work only if the number of qubits is large, in principle even infinite. Here we will show it is possible to construct different types of error correction protocols, which exploit the physical properties of the systems involved.

2 Photonic Channel

Specific descriptions of both atoms and cavities have been given in [9] and [14]. Here we give a description of transmission of qubits using photons from a more general point of view, which will allow us to optimize encoding qubits in photons. We first will define the “photonic” channel [16], and subsequently show how to establish a distant EPR pair over this channel. Once one has produced such a pair, one can transmit any qubit of information by using teleportation [11].

The standard way of encoding qubits in photons is to use two orthogonal polarization states, denoted by $+$ and $-$. Since channels are not noiseless, a polarized photon may arrive with that polarization, but also may change polarization, and/or the state may undergo phase shifts. For example, the *depolarizing channel* [17] assumes that with probability F the qubit is left intact and with probabilities $(1 - F)/3$ it undergoes a sign flip, a spin flip, or both. These errors are represented by the Pauli spin operators $\sigma_{z,x,y}$ acting on the qubit and are usually considered the most general type of error occurring. However, photons also may be absorbed, and new photons may be created. In order to include the latter errors we adopt a general formalism to describe transmitting photon states over an arbitrary channel, whose noise can be fully described by its interaction with the environment $|E\rangle$. Thus, the most general transformation of input states sent over a noisy channel is

$$\begin{aligned} |1_+\rangle|0_-\rangle|E\rangle &\longmapsto (|1_+\rangle|0_-\rangle\mathcal{T}_1 + |0_+\rangle|0_-\rangle\mathcal{T}_2 + |1_+\rangle|1_-\rangle\mathcal{T}_3 + |2_+\rangle|0_-\rangle\mathcal{T}_4 + \dots)|E\rangle, \\ |0_+\rangle|1_-\rangle|E\rangle &\longmapsto (|0_+\rangle|1_-\rangle\mathcal{T}'_1 + |0_+\rangle|0_-\rangle\mathcal{T}'_2 + |1_+\rangle|1_-\rangle\mathcal{T}'_3 + |0_+\rangle|2_-\rangle\mathcal{T}'_4 + \dots)|E\rangle, \end{aligned} \tag{1}$$

with $|n_{\pm}\rangle$ denoting a state with n photons of polarization \pm , while all other photonic modes are included in $|E\rangle$. Here the errors are written in terms of operators \mathcal{T} or \mathcal{T}' acting on the environment $|E\rangle$, where all operators are in general different, only restricted by the fact that the total transformation should be unitary. In practice, photon absorption is in fact the dominant error over longer distances. On the other hand, creation of photons in the optical regime is negligible even at room temperature. This indicates that the best way of encoding qubits into photons, is to let $|0\rangle$ correspond to sending no photons, with the simple idea that if one sends no photons, they cannot be absorbed. On the other hand, $|1\rangle$ would then correspond to one photon with particular polarization and frequency content. Thus, neglecting the creation of photons and using this encoding, we can then write for the general action of a noisy channel

$$\begin{aligned} |0\rangle &\longmapsto |0\rangle\mathcal{T}_0, \\ |1\rangle &\longmapsto |1\rangle\mathcal{T}_1 + |0\rangle\mathcal{T}_a \end{aligned} \quad (2)$$

where we have omitted the initial state of the environment. The operator \mathcal{T}_a describes the disappearance of a photon of the chosen polarization, either due to photon absorption or due to a polarization change. We emphasize that this formulation of encoding and transmission (2) incorporates more physical processes (i.e. is more general) but still is simpler than the usual one using two polarizations.

Having determined from simple physical arguments the best way of encoding photons, we now need to include in this description how (optical) photons are produced. In general, we can assume that the photon is produced and detected by making an atom change its internal state. Again, we wish to describe this process in the most general fashion. We consider two atoms A and B belonging to Alice and Bob, who have in their labs nodes 1 and 2, respectively. We denote by $|0\rangle$ and $|1\rangle$ two internal (ground-state) levels of the atoms, and by $|x\rangle$ any other level that may be involved in the process. We assume that the transmission process is such that the sending atom will produce a photon only if it started in the state $|1\rangle$, whereas no photon is produced if it is in $|0\rangle$. The receiving atom is prepared in state $|0\rangle$, and, ideally, is transferred to state $|1\rangle$ upon absorbing this photon in the correct mode, whereas if no photon arrives, the atom will remain in $|0\rangle$. Of course, errors may occur in this process: when producing a photon the first atom may undergo a transition to any other state $|x\rangle$, and similarly, the photon arriving at the second atom may induce a transition to any other state. In order to remain in the 2-dimensional Hilbert space of the atoms after the transmission we optically pump the sending atom to the state $|0\rangle$ and in the receiving atom we pump any state $|x \neq 0, 1\rangle$ to the state $|0\rangle$. Thus, the atoms undergo the following general process:

$$\begin{aligned} |\chi_0\rangle|0\rangle_A|0\rangle_B &\longmapsto |\chi_0\rangle|0\rangle_A|0\rangle_B\mathcal{T}_0, \\ |\chi_1\rangle|1\rangle_A|0\rangle_B &\longmapsto |\chi_1\rangle|0\rangle_A(|1\rangle_B\mathcal{T}_1 + |0\rangle_B\mathcal{T}_a). \end{aligned} \quad (3)$$

Here, the operators $\mathcal{T}_{0,1,a}$ contain spontaneous emission, photon absorption and transitions to other states, followed by repumping to $|0\rangle$. Thus, all complicated

physics is hidden in the three operators. We have explicitly excluded the states of all the other atoms in the network from the state $|E\rangle$ of the environment, and have included them in $|\chi_{0,1}\rangle$. In the following we will call the channel defined by (3) the “photonic channel” [16]. It describes in a general fashion the effect on atoms that are used to produce and detect photons, which in turn are transmitted over a noisy and lossy channel.

The goal is, using the photonic channel, to establish a perfect EPR pair. Since, in classical language, there are nonzero probabilities of errors described by the operators σ_z and $\sigma_- = (\sigma_x - i\sigma_y)/2$, straightforward application of the standard purification schemes would still require in principle infinitely many atoms to purify to a perfect EPR pair. Since we are interested in purification using only a few atoms, we need a different idea. The protocol designed in [14] indicated already that the restriction of using few qubits does not in principle prevent error correction schemes to be developed. There we exploited a certain property of the interaction with the environment (namely, that it is Markovian), but here we do not impose such conditions on the errors. The protocol is consequently more complicated, and we split the description in two parts. First we show how to reduce the action of the channel to that of a simpler one. Then using this simpler effective channel, we discuss the actual purification scheme.

2.1 Channel Reduction

Here we show how one can effectively eliminate the absorption term with \mathcal{T}_a from the channel (3), by using a few auxiliary atoms in each node, and by performing local operations. It is important to note here that we do not require any assumption on the operator \mathcal{T}_a .

The goal is first to reduce the channel description to

$$\begin{aligned} |\chi_0\rangle|0\rangle_A|0\rangle_B &\mapsto |\chi_0\rangle|0\rangle_A|0\rangle_B\mathcal{S}_0, \\ |\chi_1\rangle|1\rangle_A|0\rangle_B &\mapsto |\chi_1\rangle|1\rangle_A|1\rangle_B\mathcal{S}_1, \end{aligned} \quad (4)$$

with modified operators \mathcal{S} , to be determined below. In fact, we will show that, starting from a state $(|\chi_0\rangle|0\rangle_A + |\chi_1\rangle|1\rangle_A)|0\rangle_B$ we obtain either the state

$$|\chi_0\rangle|0\rangle_A|0\rangle_B\mathcal{S}_0 + |\chi_1\rangle|1\rangle_A|1\rangle_B\mathcal{S}_1, \quad (5)$$

or

$$(|\chi_0\rangle|0\rangle_A + |\chi_1\rangle|1\rangle_A)|0\rangle_B\mathcal{S}_a, \quad (6)$$

depending on the result of a set of measurements done on auxiliary atoms. In the latter case, we still have the same initial state as the environment factors out, so that we may start the procedure again, until the measurements indicate we have succeeded in obtaining the state (5). In that way, we effectively have reduced the channel to (4).

The procedure is as follows: we need 2 auxiliary atoms in node 1, denoted by A_2 and A_3 , and one in node 2, atom B_2 . First we entangle atoms A and A_2 (which starts in $|0\rangle$) according to

$$\begin{aligned}
|0\rangle_A |0\rangle_{A_2} &\mapsto |0\rangle_A |0\rangle_{A_2} \\
|1\rangle_A |0\rangle_{A_2} &\mapsto |1\rangle_A |1\rangle_{A_2}.
\end{aligned} \tag{7}$$

Then we transmit the qubit A_2 to atom B according to (3). Clearly, atom A_2 ends up in $|0\rangle$ and plays in fact the role of a dummy qubit. Thus, leaving out the state of A_2 , we obtain with this procedure the mapping

$$(|\chi_0\rangle|0\rangle_A + |\chi_1\rangle|1\rangle_A)|0\rangle_B \mapsto [|\chi_0\rangle|0\rangle_A|0\rangle_B \mathcal{T}_0 + |\chi_1\rangle|1\rangle_A(|1\rangle_B \mathcal{T}_1 + |0\rangle_B \mathcal{T}_a)], \tag{8}$$

Next we apply a local transformation between A and A_3 :

$$\begin{aligned}
|0\rangle_A |0\rangle_{A_3} &\mapsto |0\rangle_A |0\rangle_{A_3} + |1\rangle_A |1\rangle_{A_3}, \\
|1\rangle_A |0\rangle_{A_3} &\mapsto |0\rangle_A |1\rangle_{A_3} + |1\rangle_A |0\rangle_{A_3}
\end{aligned} \tag{9}$$

and proceed as follows: (i): At time t we transmit the qubit A to B , according to (8), i.e. by using the dummy atom A_2 . (ii): We interchange $|0\rangle_A \leftrightarrow |1\rangle_A$ (iii): At time $t' > t$ we transmit the qubit A to B_2 according to (8); (iv): We undo step (ii), i.e. interchange $|0\rangle_A \leftrightarrow |1\rangle_A$. After these 4 steps, we obtain the state

$$\begin{aligned}
&|0\rangle_{A_3} |\chi_0\rangle |0\rangle_A \otimes [|0\rangle_B |1\rangle_{B_2} \mathcal{T}_1(t') \mathcal{T}_0(t) + |0\rangle_B |0\rangle_{B_2} \mathcal{T}_a(t') \mathcal{T}_0(t)] \\
&+ |0\rangle_{A_3} |\chi_1\rangle |1\rangle_A \otimes [|1\rangle_B |0\rangle_{B_2} \mathcal{T}_0(t') \mathcal{T}_1(t) + |0\rangle_B |0\rangle_{B_2} \mathcal{T}_0(t') \mathcal{T}_a(t)] \\
&+ |1\rangle_{A_3} |\chi_1\rangle |0\rangle_A \otimes [|0\rangle_B |1\rangle_{B_2} \mathcal{T}_1(t') \mathcal{T}_0(t) + |0\rangle_B |0\rangle_{B_2} \mathcal{T}_a(t') \mathcal{T}_0(t)] \\
&+ |1\rangle_{A_3} |\chi_0\rangle |1\rangle_A \otimes [|1\rangle_B |0\rangle_{B_2} \mathcal{T}_0(t') \mathcal{T}_1(t) + |0\rangle_B |0\rangle_{B_2} \mathcal{T}_0(t') \mathcal{T}_a(t)]
\end{aligned} \tag{10}$$

All the terms containing the photon absorption error operator \mathcal{T}_a pertain to the combination $|0\rangle_B |0\rangle_{B_2}$. Hence, a measurement is performed on atoms B and B_2 to check whether they are in that joint state; (a) If the outcome is negative, the parts containing \mathcal{T}_a are projected out. We continue then by basically eliminating atom B_2 by performing the unitary transformation

$$|0\rangle_B |1\rangle_{B_2} \mapsto |0\rangle_B |0\rangle_{B_2} \quad |1\rangle_B |0\rangle_{B_2} \mapsto |1\rangle_B |0\rangle_{B_2}.$$

Subsequently we measure the state of A_3 . If we find $|0\rangle_{A_3}$ then we have (5) with $\mathcal{S}_0 = \mathcal{T}_1(t') \mathcal{T}_0(t)$ and $\mathcal{S}_1 = \mathcal{T}_0(t') \mathcal{T}_1(t)$, and if we find $|1\rangle_{A_3}$, we have to interchange $|0\rangle_A \leftrightarrow |1\rangle_A$, and thus obtain (5) with the definitions of $\mathcal{S}_{0,1}$ interchanged. (b) If the outcome is positive, an error has occurred and we have projected onto the terms containing \mathcal{T}_a . We measure the state of A in the 0, 1 basis, and then swap the state of A_3 into A . If the outcome was $|0\rangle_A$, then one has (6) with $\mathcal{S}_a = \mathcal{T}_a(t') \mathcal{T}_0(t)$, and if we found $|1\rangle_A$ we would instead have $\mathcal{S}_a = \mathcal{T}_0(t') \mathcal{T}_a(t)$, after interchanging $|0\rangle_A \leftrightarrow |1\rangle$. This completes the proof: either we have case (a) corresponding to (5), or case (b) and (6).

2.2 Purification Protocol

If photon absorption can be modeled by a Markovian master equation and if other (design) errors are assumed to be systematic (i.e. the same in two subsequent transmissions) we in fact have a “stationary channel” which is defined

by the condition $\mathcal{T}_1(t')\mathcal{T}_0(t)|E\rangle = \mathcal{T}_0(t')\mathcal{T}_1(t)|E\rangle$. In this case the channel (4) would allow for ideal transmission without further actions, as $\mathcal{S}_0 = \mathcal{S}_1$. This is the situation discussed in [14]. However, now we are interested in the general case where the stationarity property does not apply. In particular, this will be the case where there are additional random errors and when a Markovian description of decoherence is questionable. In the following we will show how to establish distant EPR pairs using the channel (4). The only assumption we will use, is that \mathcal{S}_0 is in some sense "close" to \mathcal{S}_1 , rather than equal. In words, we only assume that the "random" part of the errors is smaller than the systematic part (the latter including the no-error part).

Alice and Bob wish to obtain a perfect EPR pair of the form

$$|R\rangle_{AB} = |0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B \quad (11)$$

which we denoted by the "right" state $|R\rangle_{AB}$, and where we omit trivial normalization factors. In the following Alice and Bob will perform a specific protocol several times, which is such that the state of particles A and B is actually a superposition of *two* Bell states, the second one denoted by "wrong"

$$|W\rangle_{AB} = |0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B \quad (12)$$

For instance, a single use of the channel (4) will produce such a state: For convenience, we define the dual basis states $|\pm\rangle = |0\rangle \pm |1\rangle$. Alice prepares her qubit A in the state $|+\rangle_A$ and Bob prepares his qubit B in $|0\rangle_B$. They use the channel (4), and subsequently both of them apply the local operation $|0\rangle \mapsto |+\rangle$, and $|1\rangle \mapsto |-\rangle$. This way they obtain the state

$$|\Psi\rangle = |R\rangle_{AB}\frac{1}{2}(\mathcal{S}_0 + \mathcal{S}_1)|E\rangle + |W\rangle_{AB}\frac{1}{2}(\mathcal{S}_0 - \mathcal{S}_1)|E\rangle. \quad (13)$$

After this, Alice and Bob will repeat the protocol described below, and in terms of the state of the environment, at the N th step they will have a state of the form

$$|\Psi^{(N)}\rangle = |R\rangle_{AB}|E_R^{(N)}\rangle + |W\rangle_{AB}|E_W^{(N)}\rangle, \quad (14)$$

where $|E_{R,W}^{(N)}\rangle$ are unnormalized states of the environment.

In order to characterize the quality of the state (13) we define its fidelity as $F_N = |||E_R^{(N)}\rangle||^2$. The goal is to increase the fidelity to unity, so that for large N the state of the system will tend to $|R\rangle$. For the following, two auxiliary atoms A_1 and B_1 are needed, in the first and second cavity, respectively. (i) The auxiliary qubit A_1 is locally entangled with the qubit A according to the transformation

$$|0\rangle_A|0\rangle_{A_1} \mapsto |0\rangle_A|+\rangle_{A_1}; \quad |1\rangle_A|0\rangle_{A_1} \mapsto |1\rangle_A|-\rangle_{A_1}. \quad (15)$$

(ii) The qubit A_1 is transmitted to the auxiliary qubit B_1 according to the effective channel (4). Then, the qubit A_1 is measured in the $|\pm\rangle_{A_1}$ basis. If the

result is $|-\rangle_{A_1}$, one applies the unitary operation $|1\rangle_{B_1} \mapsto -|1\rangle_{B_1}$. After this, one applies a controlled rotation

$$|x\rangle_B |y\rangle_{B_1} \mapsto (-1)^{xy} |x\rangle_B |y\rangle_{B_1} \quad x, y = 0, 1. \quad (16)$$

The state after these operations is

$$|\Phi^{(N)}\rangle = |R\rangle_{AB}(|0\rangle_{B_1} \mathcal{S}_0 + |1\rangle_{B_1} \mathcal{S}_1) |E_R^{(N)}\rangle + |W\rangle_{AB}(|0\rangle_{B_1} \mathcal{S}_0 - |1\rangle_{B_1} \mathcal{S}_1) |E_W^{(N)}\rangle. \quad (17)$$

(iii) The auxiliary qubit B_1 is measured. If one finds B_1 in the $|+\rangle$ state, one gets

$$|E_R^{(N+1)}\rangle = \frac{1}{2}(\mathcal{S}_0 + \mathcal{S}_1) |E_R^{(N)}\rangle, \quad (18)$$

$$|E_W^{(N+1)}\rangle = \frac{1}{2}(\mathcal{S}_0 - \mathcal{S}_1) |E_W^{(N)}\rangle, \quad (19)$$

In this case, the “right” state is multiplied by the operator $\mathcal{S}_0 + \mathcal{S}_1$, which is a “larger” operator than the $\mathcal{S}_0 - \mathcal{S}_1$ combination, which multiplies the “wrong” state. Thus, in this case the fidelity increases, in a manner specified below. If, on the other hand, the result of the measurement on B_1 is $|-\rangle$, one has

$$|E_R^{(N+1)}\rangle = \frac{1}{2}(\mathcal{S}_0 - \mathcal{S}_1) |E_R^{(N)}\rangle, \quad (20)$$

$$|E_W^{(N+1)}\rangle = \frac{1}{2}(\mathcal{S}_0 + \mathcal{S}_1) |E_W^{(N)}\rangle, \quad (21)$$

which implies that the fidelity has decreased. We denote by P_{\pm} the probability of these two outcomes.

2.3 Analysis of Protocol

We analyze how the fidelity changes after each step, for which we need to evaluate P_{\pm} . To this end we define for arbitrary states of the environment the quantities

$$\pi_{\pm} \equiv \frac{\|\frac{1}{2}(\mathcal{S}_0 \pm \mathcal{S}_1)|E\rangle\|^2}{\| |E\rangle \|^2}. \quad (22)$$

To calculate π_{\pm} one in principle would need to know the specific form of the operators and states at all times. We assume for simplicity, however, that the π_{\pm} do not depend on $|E\rangle$. Moreover, the only assumption on the operators $\mathcal{S}_0 \pm \mathcal{S}_1$ is that $\pi_+ > \pi_-$. Using the definition (22) we have then

$$P_{\pm} = \pi_{\pm} F_N + \pi_{\mp} (1 - F_N), \quad (23)$$

which implies that $P_+ > P_-$. Hence, the probability that the fidelity increases is larger than the probability that F decreases. In fact, corresponding to the measurement $|\pm\rangle$ the new fidelity is

$$F_{N+1} = \frac{\pi_{\pm} F_N}{\pi_{\pm} F_N + \pi_{\mp} (1 - F_N)}, \quad (24)$$

respectively. For $\pi_+ > \pi_-$, the outcome $|+\rangle_{B_1}$ increases the fidelity and occurs with a higher probability. Since the decrease in fidelity after a $|-\rangle_{B_1}$ measurement is compensated for by a subsequent $|+\rangle_{B_1}$ measurement, the protocol consists of a random walk along a set of particular values of F , where it is more likely to go up than to go down, thus achieving $F \uparrow 1$ asymptotically. The effectiveness of the process depends only on the value of π_+ which characterizes the effective channel. We have simulated the improvement of the fidelity for several values of the probability π_+ . For example, we obtain $F = 1 - 10^{-5}$ after $N = 12$ steps for $\pi_+ = 0.9$, which indicates that the scheme is quite effective. In fact, unit fidelity is approached exponentially fast with the number of steps. This can be shown analytically as follows: First, we note that, if the number of upward steps in the random walk in excess of downward steps equals N_1 , then the fidelity is

$$F(N_1) = \frac{N_1}{1 + \left(\frac{\pi_-}{\pi_+}\right)^{N_1}} \quad (25)$$

so that, conversely, in order to reach a final fidelity $F_\infty \equiv 1 - \delta$ one needs approximately

$$N_1(\delta) = \frac{\log\left(\frac{\delta}{1-\delta}\right)}{\log(\pi_-/\pi_+)} \quad (26)$$

steps. Second, since not every measurement will have the outcome $+$, we also need the number of steps N_2 needed to reach N_1 steps up in excess of steps down. This is approximated by the average of $N_1/(P_+ - P_-)$,

$$N_2 \approx \int_{\pi_+}^{1-\delta} dF \frac{N_1}{S(2F-1)} = \frac{N_1}{2S} \log \frac{1-2\delta}{S}, \quad (27)$$

where $S = \pi_+ - \pi_-$, and where we used that the initial value of F equals π_+ . The total number of steps required to reach a fidelity $1 - \delta$ is, therefore,

$$N_2 \approx \frac{1}{2S} \log \frac{1-2\delta}{S} \frac{\log\left(\frac{\delta}{1-\delta}\right)}{\log(\pi_-/\pi_+)} \approx \frac{\log(1/S) \log(1/\delta)}{2S \log(\pi_+/\pi_-)} \quad (\delta \rightarrow 0) \quad (28)$$

which for small δ behaves as $\log(1/\delta)$, so that conversely the fidelity grows exponentially fast to 1 with the number of steps taken.

Finally, let us note that for a “good” standard channel [see Eq. (3)], \mathcal{T}_0 is close to \mathcal{T}_1 . As a consequence of our reduction procedure this implies that $\mathcal{S}_0 \simeq \mathcal{S}_1$ and therefore $\pi_+ \simeq 1$, and $\pi_- \ll 1$. We emphasize that the reverse statement is not true, namely one can have $\mathcal{S}_0 \simeq \mathcal{S}_1$ but a “bad” standard channel. This shows our procedure introduces a new element, allowing one in certain cases to develop improved error correction schemes.

3 Further Developments

So far we tacitly assumed all local operations to be perfect, which is the standard assumption for designing purification protocols. The reason is that nonlocal

operations are, of course, much more difficult to control than local operations. However, in practice also 2-bit local operations are not at all easy to perform. For instance, only very recently has one succeeded in entangling two atoms [4]. Since on the other hand, 1-bit operations can be considered routine, an important question to answer is, whether one can "purify" 2-bit local operations provided 1-bit operations are perfect. This question we have answered recently in [18], where it was shown that, e.g., the joint measurement operation needed for the purification protocol discussed above can actually be purified. A crucial property is that this measurement can be performed "fault-tolerantly", in the sense that an error occurring during the measurement is always detected by the measurement itself. Moreover, with a more complicated procedure, also a universal 2-bit gated can be performed perfectly in the same sense that errors may occur, but will always be detected.

A second avenue of further research is the following: the error correction protocol presented above works to all orders in the photon absorption error. The larger the probability for this error to occur is, the more often one has to go through all the steps of the protocol. Since the error probability scales exponentially with the fiber length L , the number of tries in the protocol, and hence the time required for sending a qubit, increases exponentially with L as well. In order to avoid this, we can divide the fiber into N segments and place "quantum repeaters" at each segment, just as in classical communication. An analysis of this problem [19] shows that the fidelity of EPR pairs thus obtained decreases exponentially with the number of segments. This drawback, however, is not very serious, as above-described error correction scheme can increase the fidelity exponentially fast with the number of steps, as just shown. On the other hand, including now also local errors, this is no longer true, and another higher-level purification scheme is needed. This is reminiscent of concatenated quantum codes for fault-tolerant quantum computation [13]. A main difference is that in the present case it is not required to be able to protect arbitrary states sent over a channel, as we only attempt to establish an EPR pair. As a result, the threshold for tolerable errors in local operations is a few percent, rather than 10^{-5} .

Acknowledgments

This work was supported in part by the Österreichischer Fonds zur Förderung der wissenschaftlichen Forschung, by the European TMR network ERB4061PL95-1412, by NSF PHY94-07194 and PHY-93-13668, by DARPA/ARO through the QUIC program, and by the ONR.

References

1. Proceedings of the Nobel Symposium 104, 13-17 June 1997, to be published in *Physica Scripta*.

2. C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano, and D.J. Wineland, Phys. Rev. Lett. **75**, 4714 (1995).
3. Q.A.Turchette, C.J.Hood, W.Lange, H.Mabuchi, and H.J.Kimble, Phys. Rev. Lett. **75**, 4710 (1995).
4. E. Hagley, X. Maitre, G. Nogues, C. Wunderlich, M. Brune, J. M. Raimond, and S. Haroche, Phys. Rev. Lett. **79**, 1 (1997).
5. See, for example, D. P. DiVincenzo, Science **270**, 255 (1995).
6. J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).
7. T. Pellizzari, S. A. Gardiner, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **75**, 3788 (1995).
8. K. Mattle, H. Weinfurter, P.G. Kwiat and A. Zeilinger, Phys. Rev. Lett. **76**, 4656 (1996).
9. J. I. Cirac, P. Zoller, H. Mabuchi, and H. J. Kimble, Phys. Rev. Lett. **78**, 3221 (1997).
10. C. H. Bennett, Phys. Today **24** (October 1995) and references therein; A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
11. C. H. Bennett *et al*, Phys. Rev. Lett. **70**, 1895 (1993).
12. C. H. Bennett *et al.*, Phys. Rev. Lett. **76**, 722 (1996); D. Deutsch *et al.*, Phys. Rev. Lett. **77**, 2818 (1996); N. Gisin, Phys. Lett. A **210** 151 (1996); A. Ekert and C. Macchiavello, Phys. Rev. Lett. **77**, 2585 (1996).
13. P.W. Shor, Phys. Rev. A **52**, R2493 (1995); A.M.Steane, Phys. Rev. Lett. **77**, 793 (1996); J. I. Cirac, T.Pellizzari and P. Zoller, Science **273**, 1207 (1996); E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997); P. Shor, *Fault-tolerant quantum computation*, quant-ph/9605011; D. DiVincenzo and P.W. Shor, Phys. Rev. Lett. **77**, 3260 (1996).
14. S. J. van Enk, J. I. Cirac and P. Zoller, Phys. Rev.Lett. **78**, 4293 (1997).
15. T. Pellizzari, Phys. Rev. Lett. **79**, 5242 (1997).
16. S. J. van Enk, J. I. Cirac, and P. Zoller, Science **279**, (1998).
17. See, for example, B. Schumacher, Phys. Rev. A **45**, 2614 (1996); C. H. Bennett, D.P. DiVincenzo and J.A. Smolin, Phys. Rev.Lett. **78**, 3217 (1997).
18. S. J. van Enk, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **79**, 5178 (1997).
19. H.-J. Briegel, W. Dür, J.I. Cirac, and P. Zoller, quant-ph/9803056.

An Optical Approach to Quantum Computing

J.D. Franson and T.B. Pittman

Applied Physics Laboratory
The Johns Hopkins University
Laurel, MD 20723, USA

1 Introduction

Any realistic approach to quantum computing must be capable of implementing a sufficiently large number of qubits to perform useful computations. It has been estimated that 10^4 qubits may be required to implement Shor's factoring algorithm for integers of useful size, and that this number may grow to 10^6 qubits when the redundant bits required for quantum error correction are included. Many of the quantum computer implementations currently being investigated, such as ion traps or NMR techniques, provide a reasonably straightforward method for implementing a few qubits but are subject to fundamental limitations on the number of qubits that can be implemented [1, 2].

The long-term goal of our work is to develop an optical approach to quantum computing in which the logic gates and memory devices have a sufficiently simple structure that large numbers of devices could be mass-produced at a reasonable cost. This is a modular approach in which independent logic gates could be connected using optical fibers or waveguides. The intrinsic decoherence of the devices is expected to be sufficiently small that quantum error correction techniques may eventually be able to compensate for them. In that case, the modular nature of the approach and the simple structure of the devices may allow the implementation of a full-scale computer.

One of the main difficulties in any optical approach to quantum computing is the small magnitude of the nonlinear interaction between two photons, which is required for the implementation of quantum logic gates. Other research groups [3] have obtained nonlinear phase shifts at the two-photon level by confining the photons to high-Q cavities and using atomic beams or traps, but that approach may not be feasible for the construction of full-scale computers due to the size and complexity of the cavities and traps. Our approach relies on a newly-predicted mechanism [4] for the production of nonlinear phase shifts that is non-classical and exists only at the quantum level. This mechanism is expected to give nonlinear phase shifts that are orders of magnitude larger than those obtained previously, which should eliminate the need for resonant cavities and atomic beams or traps.

2 Nonlocal Enhancement of Nonlinear Phase Shifts

Nonlinear phase shifts at the two-photon level can be used to implement quantum logic gates, as will be discussed in more detail below. Previous mechanisms for the

production of nonlinear phase shifts give a phase shift that is proportional to the number N_A of atoms in the medium, since both photons must interact with the same atom. Our new mechanism involves the interaction of photons with pairs of atoms in the medium. Since the number of pairs of atoms is proportional to N_A^2 , this gives a nonlinear phase shift proportional to N_A^2 and a large enhancement of the phase shift for large values of N_A .

The new mechanism [4] for the production of nonlinear phase shifts is illustrated in Figure 1. Two photons with frequencies ω_1 and ω_2 propagate through a medium and interact with two atoms, such as those labeled A and B in the figure. As illustrated in the Feynman-like diagram of Figure 2, atom A absorbs photon 1 and re-emits photon 2, while atom B absorbs photon 2 and re-emits photon 1. This exchange of the two photons has no net effect other than to produce an energy shift ΔE and a corresponding phase shift that can be calculated using fourth-order perturbation theory. Each pair of atoms corresponds to a distinct Feynman diagram, so that the sum over intermediate states gives a factor of N_A^2 . Since N_A can be on the order of 10^{12} , this interaction is expected to give phase shifts that are many orders of magnitude larger than those obtained using earlier methods.

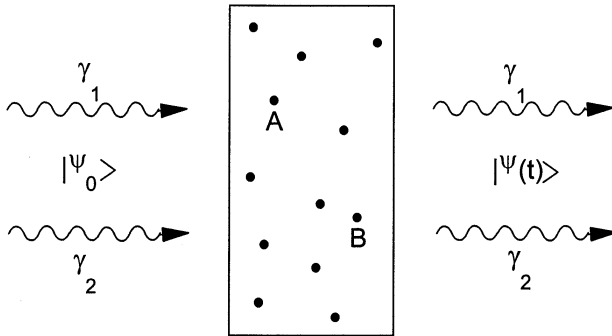


Fig. 1. Nonlinear phase shift that occurs when two photons pass through a medium containing N_A atoms, two of which are labeled A and B.

The total contribution from all Feynman diagrams of this kind gives a net phase shift proportional to

$$\Delta E = -2 \frac{M^4 N_A^2 n_1 n_2 f_R}{\delta^3} \frac{w^2}{(\delta_1 - \delta_2)^2} \quad (1)$$

Here ΔE is the energy shift calculated using perturbation theory, δ_1 and δ_2 are the photon detunings (difference in photon energy from the atomic excitation energy), M is the atomic matrix element, n_1 and n_2 are the number of photons in each beam, w is the line width due to collisions, and f_R is a dimensionless

factor on the order of unity that is associated with atomic recoil . Under the appropriate experimental conditions, it can be shown that the nonlinear phase shift is approximately related to the usual, linear phase shift (index of refraction) by

$$\Delta\phi_{non} \sim \Delta\phi_{lin}^2 \quad (2)$$

It is relatively easy to obtain linear phase shifts of π radians with negligible decoherence due to absorption and scattering, in which case the nonlinear phase shift should also be on the order of unity with very small decoherence.

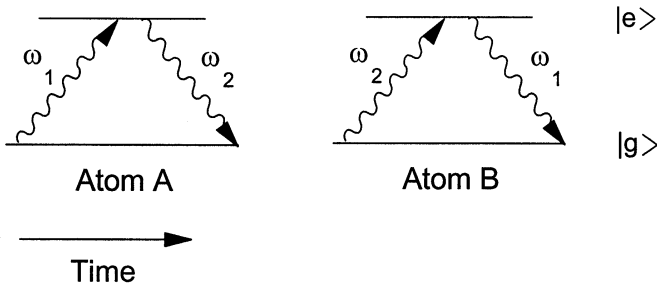


Fig. 2. Typical Feynman-like diagram in which the exchange of two photons by a pair of atoms gives a nonlinear phase shift proportional to N_A^2 .

Since the two photons have different energies, the process shown in Figure 2 does not conserve energy at the location of atom A nor at the location of atom B, but it does conserve energy overall. Roughly speaking, each atom must “know” of the presence of the other atom in order to conserve the overall energy at the end of the process, which suggests that this is a nonlocal process somewhat similar to the EPR paradox. It can be shown that this effect cannot be derived using the nonlinear susceptibility coefficients commonly employed in nonlinear optics. In addition, phase shifts of this kind only occur for states with a well-defined number of photons and no such effect is expected to occur for classical beams of light.

3 Preliminary Experimental Results

A preliminary series of experiments have been performed in a sodium vapor cell using the apparatus outlined in Figure 3. Without describing any of the details, the experiment was configured in such a way that the nonlinear phase shifts described in section 2 would cause the presence of photon 1 to rotate the linear polarization of photon 2 if both photons passed through the sodium vapor cell at the same time. A polarization analyzer oriented at 90° to the original polarization

of photon 2 was placed in its path after it had passed through the sodium vapor cell. If photon 1 passed through the cell at the same time, the rotation of the polarization of photon 2 would allow it to pass through the analyzer and be detected. If the two photons passed through the cell at different times, no such polarization rotation would occur and photon 2 would not be detected (for a perfect analyzer). Thus the presence or absence of photon 1 could control the path taken by photon 2, which corresponds to an elementary logic switch.

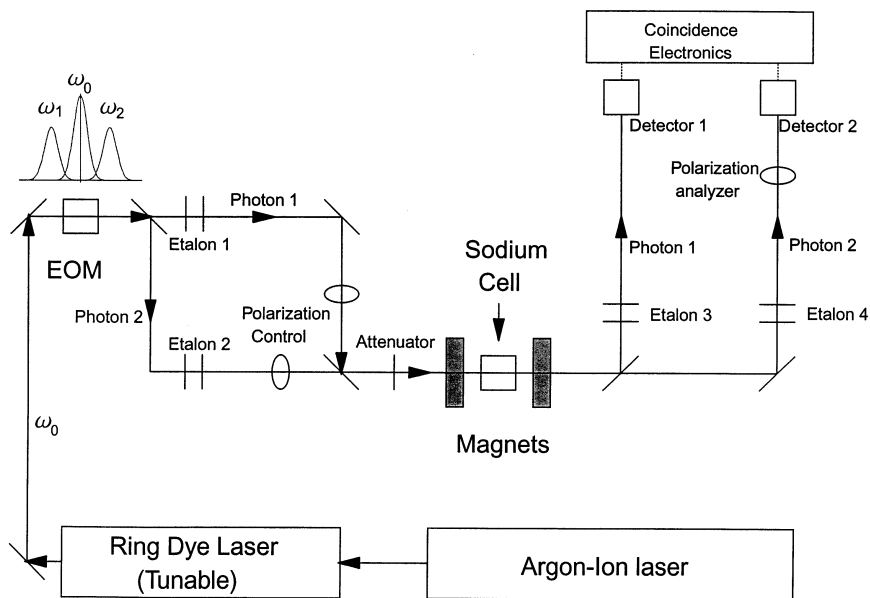


Fig. 3. Simplified design of the experimental apparatus.

An effect of this kind can be seen in the preliminary experimental results of Figure 4, which is a plot of the rate at which both photons were detected as a function of the difference in their arrival times. The peak at zero time delay suggests that photon 1 can control the polarization of photon 2, as expected. This peak is significant at the three-standard-deviation level, based on a careful Monte Carlo simulation of the data collection process and the observed data. The probability that such a peak is a statistical anomaly is approximately 10^{-3} . Smaller peaks were observed at lower sodium densities as expected.

Dispersion relations based on causality provide a connection between the real part of the index of refraction (phase shifts) and the imaginary part (absorption or scattering). The imaginary part of the index of refraction was also measured by removing the polarization analyzer and focusing the photon 2 beam near the

edge of the corresponding detector, so that any scattering would be expected to produce a dip in the coincidence curve rather than a peak. The results of measurements of that kind show a dip in the coincidence curve at zero time delay, as expected.

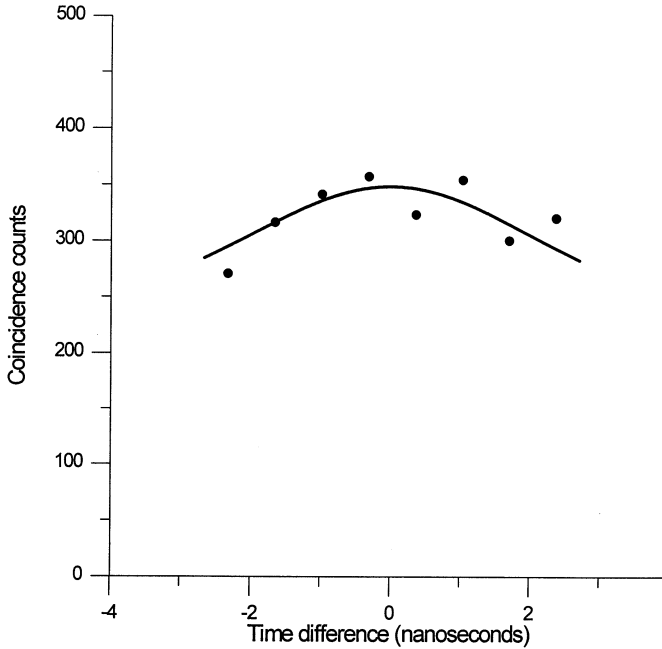


Fig. 4. Preliminary results of nonlinear phase-shift measurements.

At the time of this conference, similar effects have been seen in a number of sodium vapor cells. These effects vanish at low sodium vapor pressures and in the absence of any buffer gas, as would be expected from the theory. We are now preparing to make a series of systematic measurements in order to verify the dependence of the phase shift on N_A^2 . This will require a new design for the sodium vapor cell and the use of higher speed electronics. Although systematic measurements of this kind are required in order to conclusively demonstrate the existence of the predicted effect, the preliminary results do suggest that nonlinear phase shifts can be obtained at the two-photon level without the use of resonant cavities.

Similar effects are expected in crystalline materials, which provide higher atomic densities and long-term stability. The use of solid-state materials will probably be necessary for the construction of useful quantum logic gates. An investigation of these nonlinear phase shifts in various crystalline materials is planned.

4 Quantum Computer Implementation

We have also investigated the implementation of a quantum computer using these nonlinear effects in some detail. The implementation of a Controlled-NOT gate is shown in Figure 5. Here the nonlinear phase shift is employed in an interferometer arrangement to allow photon 1 to control the output path taken by photon 2 [5]. Additional nonlinear phase shifts have been included to provide the conventional overall phase shifts for each of the possible input states.

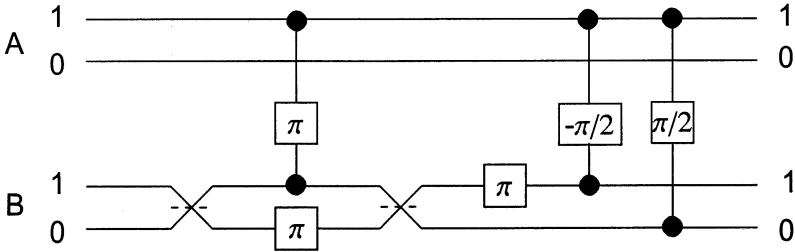


Fig. 5. Implementation of a Controlled-NOT gate giving the conventional phase shifts.

The new mechanism for nonlinear phase shifts requires that the two photons have different frequencies ω_1 and ω_2 . The implementation of a practical quantum computer would, instead, require logic gates acting on two photons of the same frequency. This can be accomplished by introducing a scratch bit corresponding to a photon at a different frequency, as illustrated in Figure 6. The scratch bit is returned to its original state at the end of the operation, so that “garbage” bits do not accumulate.

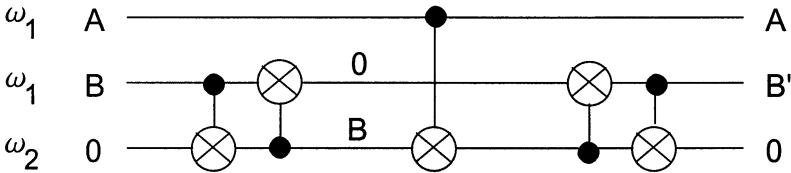


Fig. 6. A Controlled-NOT gate acting on two photons A and B with the same frequency. A scratch qubit at a different frequency is used but returned to its initial state of 0.

In this approach, the memory registers would be implemented by switching a single photon into one of two optical-fiber loops, as illustrated in Figure 7. We have estimated that several million logic operations could be performed during the time that a photon could be stored in this way. Quantum error correction techniques could then be used to extend the effective storage time to much larger values. Several techniques for compensating for the effects of dispersion are currently under consideration.

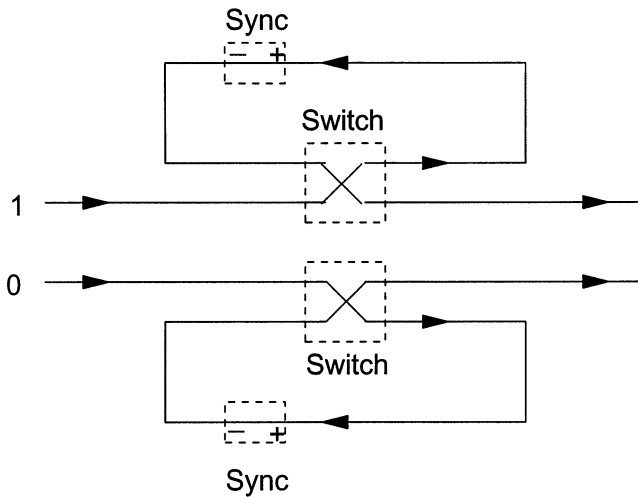


Fig. 7. Proposed memory storage device consisting of two loops of optical fiber with electro-optical switches.

5 Summary

An optical approach to quantum computing is expected to have many potential advantages, including the ability to connect independent logic gates with optical fibers or waveguides. We are investigating a new mechanism for the production of the required nonlinear phase shifts at the two-photons level, which is based on nonlocal interactions between pairs of atoms in a medium. A preliminary series of experiments provides some evidence for the existence of the predicted nonlinear phase shifts. The implementation of a quantum computer using this mechanism has been considered in some detail.

This approach relies on new physics that has not been thoroughly investigated, and unexpected difficulties may be encountered. Nevertheless, the modular nature of the approach, the simplicity of the logic devices, and its relatively low rate of decoherence may eventually allow the construction of full-scale quantum computers.

References

1. Warren, W. (1997). The Usefulness of NMR Quantum Computing. *Science*, **277**, 1688–1689.
2. Hughes, R.J. (1997). Decoherence Bounds to Quantum Computing with Trapped Ions. *Proceedings of the Workshop on Fundamental Problems in Quantum Theory*, Baltimore, Aug. 3-7, 1997. To be published in *Fortschritte der Physik*.
3. Turchette, Q.A., Hood, C.J., Lange, W., Mabuchi, H. and Kimble, H.J. (1995). Measurement of Conditional Phase Shifts for Quantum Logic. *Phys. Rev. Lett.*, **75**, 4710–4713.
4. Franson, J.D. (1997). Cooperative Enhancement of Optical Quantum Gates. *Phys. Rev. Lett.*, **78**, 3852–3855.
5. Milburn, G.J. (1989). Quantum Optical Fredkin Gates. *Phys. Rev. Lett.*, **62**, 2124–2127.

Quantum Computation with Linear Optics

C. Adami and N.J. Cerf

W. K. Kellogg Radiation Laboratory
California Institute of Technology, Pasadena, California 91125, USA

Abstract. We present a constructive method to translate small quantum circuits into their optical analogues, using linear components of present-day quantum optics technology only. These optical circuits perform precisely the computation that the quantum circuits are designed for, and can thus be used to test the performance of quantum algorithms. The method relies on the representation of several quantum bits by a single photon, and on the implementation of universal quantum gates using simple optical components (beam splitters, phase shifters, etc.). The optical implementation of Brassard et al.'s teleportation circuit, a non-trivial 3-bit quantum computation, is presented as an illustration.

1 Introduction

The promise of ultrafast computation using quantum mechanical logic raised by Shor's discovery of a polynomial algorithm for factoring [1] has yet to materialize in physical implementations. While quantum logic has been implemented in a number of different guises [23], the dynamics and behavior of a quantum circuit subject to noise and quantum decoherence has only been tested in simulations on a classical computer [45] (but see [6]). There is little controversy about the realization that it is the quantum mechanical superposition principle, and the entangled, nonlocal, states it engenders, that are at the origin of the speed-up of quantum algorithms with respect to their classical counterparts. Still, effective quantum algorithms are few and far between, and even those that are known today have yet to be tested in a physical realization (but see [7].)

In anticipation of physical realizations that implement quantum superpositions between physical states, we present here a method of constructing circuits based on non-local superpositions of "eventualities", rather than physical objects. More precisely, we *simulate* quantum superpositions, "qubits", as "which-path" eventualities in linear optics, implementable on standard optical benches. While the "support" of these qubits is decidedly classical (the optical devices such as beam splitters, polarizers, etc.) the wave function at the exit of the optical circuit can be made to coincide arbitrarily well with the outcome of the anticipated computation, thus implementing the quantum circuit. Naturally, this "classical" implementation of quantum logic has its drawbacks, as we comment on further below. Still, it should provide an excellent (and cost effective) means for testing small circuits for quantum error correction or quantum algorithms.

As we point out below, the realization that optical which-path eventualities of single photons can simulate qubits is not new in itself. Here, we focus on *protocols* to translate *any* quantum circuit diagram into *linear* optics networks, which puts the realization of simple circuits decidedly within reach. Quantum computation can be described as the task of performing a specific unitary transformation on a set of quantum bits (qubits) followed by measurement, so that the outcome of the measurement provides the result of the computation. This unitary transformation can be constructed with a finite number of 4×4 unitary matrices, that is, using a quantum circuit utilizing only 1-bit and 2-bit quantum gates (see, *e.g.*, [89]). The universality of 1- and 2-bit gates in the realization of an arbitrary quantum computation was shown in [10]. Furthermore, it was realized recently that an *optical* realization exists for any $N \times N$ unitary matrix [11], a result which generalizes the well-known implementation of $U(2)$ matrices using a lossless beam splitter and a phase shifter (see, *e.g.*, [12]). Accordingly, each element of $U(N)$ can be constructed using an array of $\mathcal{O}(N^2)$ beam splitters that form an optical multiport with N input and N output beams. As we shall see below, this result together with the universality of (1- and 2-bit) gates, can be exploited *constructively*, providing a systematic method for assembling optically-simulated gates to build simple quantum circuits.

2 Logical Qubits in Optics

Let us start by considering the equivalence between traditional linear optics elements (such as beam splitters or phase shifters) and 1-bit quantum gates (see, *e.g.*, [13]). This equivalence is inspired by the standard two-slit experiment of quantum mechanics, in which a single quantum can interfere with itself to produce fringes on a screen. Accordingly, a quantum on the other side of the slit is in a superposition of paths, and the quantum mechanical uncertainty principle is in full effect with respect to location and phase [14].

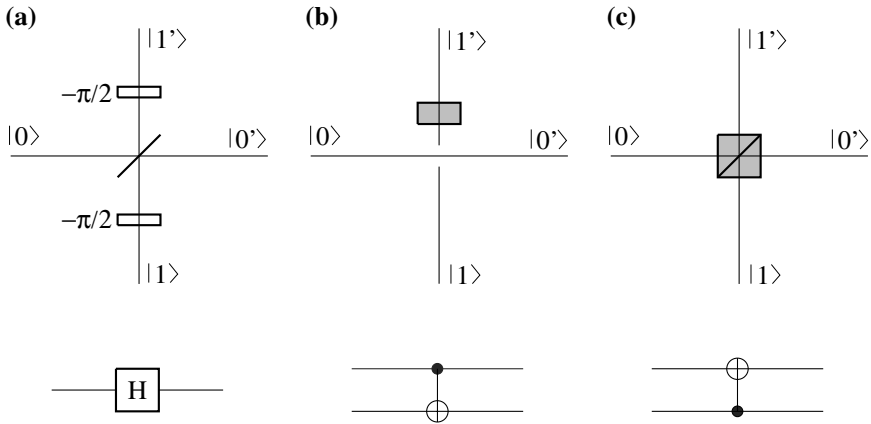
For example, in quantum circuit terminology, an optical symmetric beam splitter is known to act as a quantum $\sqrt{\text{NOT}}$ gate (up to a phase of $\pi/4$) if we use the pair of input modes $|01\rangle$ (or $|10\rangle$) to represent the logical 0 (or 1) state of the qubit. If one input port is in the vacuum state $|0\rangle$ (absence of a photon) and the second one is in a single-photon state $|1\rangle$, the output ports will then be in a superposition state $|01\rangle + i|10\rangle$. Thus, with the identification of the *logical* qubits $|0_L\rangle \equiv |01\rangle$ and $|1_L\rangle \equiv |10\rangle$, we produce the wavefunction

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|0_L\rangle + i|1_L\rangle) , \quad (1)$$

from the initial state $|0_L\rangle$ just by running a photon through a beam splitter. (The factor i arises from the $\pi/2$ phase shift between the transmitted and the reflected wave in a lossless symmetric beam splitter [?].)

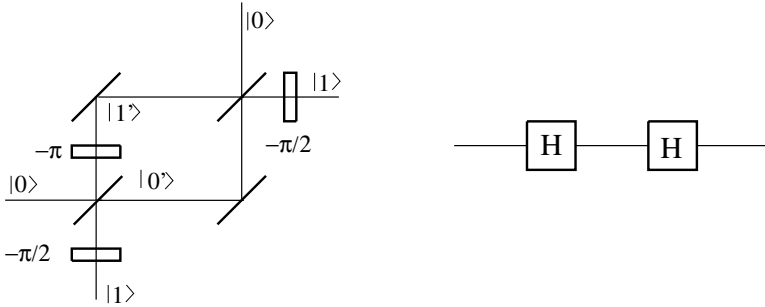
Similarly, a quantum phase gate can be obtained by use of a phase shifter acting on one mode of the photon. In other words, single-photon interferometry experiments can be interpreted in quantum circuit language, the “which-path”

Fig. 1. Example of optical simulation of basic quantum logic gates. (a) Hadamard gate on a “location” qubit, using a lossless symmetric beam splitter. (b) Controlled-NOT gate using a polarization rotator. The location and polarization are the control and target qubit, respectively. (c) Same as (b) but the control and target qubits are interchanged by the use of a polarizing beam splitter.



variable being substituted with a quantum bit. Although a general proof for the existence of an optical realization of an arbitrary quantum circuit is implicitly given in Ref. [11], the simple duality between quantum logic and single-photon optical experiments has not been exploited. Here (and in [16]) it is shown that a *single-photon* representation of *several* qubits can be used to exploit this duality: as several (say n) quantum bits can be represented by a single photon in an interferometric setup involving essentially 2^n paths, quantum conditional dynamics can easily be implemented by using different optical elements in distinct paths. The appropriate cascading of beam splitters and other linear optical devices entails the possibility of simulating networks of 1- and 2-bit quantum gates (such as the Hadamard or the controlled-NOT gate, see Fig. 1), and thereby in principle achieving universal n -bit quantum computations [16]. This is in contrast with traditional optical models of quantum logic, where in general n photons interacting through *nonlinear* devices (acting as 2-bit quantum gates) are required to represent n qubits (see, e.g., [13]). Such models typically make use of the Kerr nonlinearity to produce intensity-dependent phase shifts, so that the presence of a photon in one path induces a phase shift to a second photon (see, e.g., the optical realization of a Fredkin gate [17]). Instead, the method proposed here yields a straightforward method for “translating” *any* n -bit quantum circuit into a single-photon optical setup, whenever n is not too large. The price to pay is the exponential growth of the number of optical paths, and, consequently, of

Fig. 2. Implementation of two sequential Hadamard transformations as a balanced Mach-Zehnder interferometer using lossless symmetric beam splitters only.



optical devices that are required. This will most likely limit the applicability of the proposed technique to the implementation of relatively simple circuits.

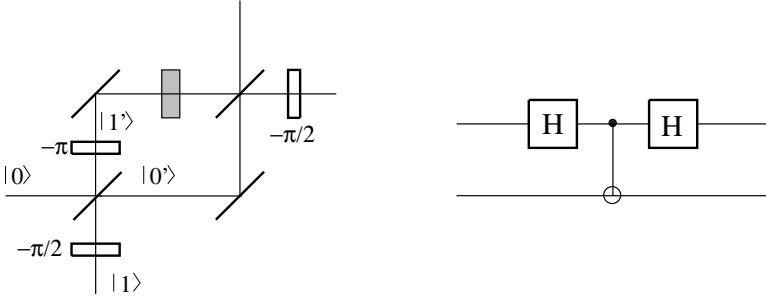
First, let us consider a single-photon experiment with a Mach-Zehnder interferometer in order to illustrate the optical simulation of elementary quantum gates (see Fig. 1). One qubit is involved in the description of the interferometer in terms of a quantum circuit: the “location” qubit, characterizing the information about “which path” is taken by the photon. Rather than using the occupation number representation for the photon, here we label the two input modes entering the beam splitter by $|0\rangle$ and $|1\rangle$ (“mode description” representation). The quantum state of the photon *exiting* the beam splitter then is $|0'\rangle + i|1'\rangle$ or $|1'\rangle + i|0'\rangle$ depending on the input mode of the photon. This is the $\sqrt{\text{NOT}}$ gate discussed earlier. Placing phase shifters at the input and output ports as shown in Fig. 1a, the beam splitter can be shown to perform a Hadamard transformation between input and output modes, *i.e.*,

$$\begin{pmatrix} |0'\rangle \\ |1'\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} |0\rangle \\ |1\rangle \end{pmatrix}. \quad (2)$$

In this sense, a lossless symmetric beam splitter (supplemented with two $-\pi/2$ phase shifters) can be viewed as a Hadamard gate acting on a location qubit. Recombining the two beams using a second beam splitter (see Fig. 2) in order to form a balanced Mach-Zehnder interferometer corresponds therefore, in this quantum circuit language, to having a second Hadamard gate acting subsequently on the qubit¹. Since $H^2 = 1$, it is not a surprise that the location qubit returns to the initial basis state ($|0\rangle$ or $|1\rangle$) after two beam splitters (with the appropriate phase shifter). This sequence of two Hadamard gates simply conveys the fact that the contributions of the two paths interfere destructively in

¹ Here and below, it is understood that the path lengths are adjusted so that the difference between dynamical phases vanishes.

Fig. 3. Implementation of two sequential Hadamard transformations with intermediate *conditional* operation on the polarization. This circuit produces an entangled state $|0\rangle_{\text{pol}}|0\rangle_{\text{loc}} + |1\rangle_{\text{pol}}|1\rangle_{\text{loc}}$ between the polarization (which is initially in a product state with the location qubit at the input port of the interferometer) and the location qubit (denoted by $|0\rangle$ and $|1\rangle$ in the figure) just before the final beam splitter, preventing the observation of interference fringes at the output of this interferometer.



one of the output ports, so that the photon always leaves the interferometer in the other.

More interestingly, consider now the same interferometer using polarized photons (the photon is horizontally polarized at the input). Assuming that none of the devices acts on polarization, the photon exits the interferometer with the same polarization. In a circuit terminology, this corresponds to introducing a “polarization” qubit ($|0\rangle_{\text{pol}}$ stands for horizontal polarization) which remains in a product state with the location qubit throughout the circuit. If a polarization rotator is placed in one of the branches of the interferometer, flipping the polarization from horizontal $|0\rangle_{\text{pol}}$ to vertical $|1\rangle_{\text{pol}}$, it is well known that interference disappears since both paths become distinguishable. This corresponds to placing a 2-bit controlled-NOT gate (represented in Fig. 1b) between the two Hadamard gates, where the location qubit is the control and polarization is the target bit (see Fig. 3). The circuit in Fig. 3 thus simply implements the dynamics

$$|0\rangle_{\text{pol}}|0\rangle_{\text{loc}} \rightarrow \frac{1}{\sqrt{2}} (|0\rangle_{\text{pol}}|0\rangle_{\text{loc}} + |1\rangle_{\text{pol}}|1\rangle_{\text{loc}}) \quad (3a)$$

$$\rightarrow \frac{1}{2} (|0\rangle_{\text{pol}}|0\rangle_{\text{loc}} + |1\rangle_{\text{pol}}|0\rangle_{\text{loc}} + |0\rangle_{\text{pol}}|1\rangle_{\text{loc}} - |1\rangle_{\text{pol}}|1\rangle_{\text{loc}}) . \quad (3b)$$

which “tags” each path with a particular polarization just before the final beam splitter in the sense that the polarization of the photon is flipped *conditionally* on its location. The disappearance of interference fringes then simply reflects the entanglement between location and polarization qubits (the reduced density matrix obtained by tracing over polarization shows that the photon ends up in a mixed “location” state, *i.e.*, it has a 50% chance of being detected in one or

the other exit port). This suggests that Feynman's rule of thumb (namely that interference and which-path information are complementary) is a manifestation of the quantum *no-cloning* theorem: the location qubit cannot be "cloned" into a polarization qubit. However, the fringes can be resurrected via a *quantum erasure* procedure [18] (which involves placing polarizing beam-splitters, introduced below, at the exit ports of the construction).

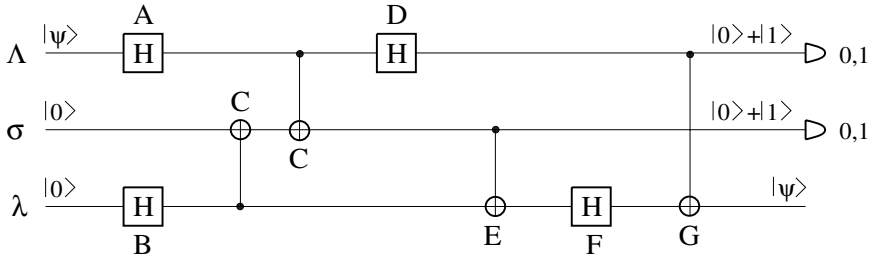
The optical analogue of other basic quantum gates can be devised following the same lines. For example, a polarizing beam splitter achieves a controlled-NOT gate where the location qubit is flipped or not (the photon is reflected or not) conditionally on its state of polarization, as shown in Fig. 11c. Fredkin, Toffoli, as well as controlled-phase gates can easily be simulated in the same manner but will not be considered here. The central point is that, in principle, a universal quantum computation can be simulated using these optical substitutes for the universal quantum gates. The optical setup is constructed straightforwardly by inspection of the quantum circuit. A circuit involving n qubits requires in general n successive splitting stages of the incoming beam, that is, 2^n optical paths are obtained via $2^n - 1$ beam splitters. (The use of polarization of the photon as a qubit allows using 2^{n-1} paths only.) This technique is thus limited to the simulation of quantum networks involving a relatively small number of qubits (say less than 5-6 with present technology). The key idea of a quantum computer, however, is to avoid just such an exponential size of the apparatus by having n physical qubits performing unitary transformations in a 2^n -dimensional space. In this respect, it can be argued that an optical setup requiring $\sim 2^n$ optical elements to perform an n -bit quantum computation represents a *classical* optical computer (see, e.g., [9]). Accordingly, the issue of whether non-locality (which is at the heart of entanglement) is *physically* present in the optical realization is a matter of debate.

3 Optical Quantum Teleportation

As an illustration, we show that a quantum circuit involving 3 qubits and 8 quantum gates (see Fig. 4) can be implemented optically using essentially 9 beam splitters [16]. This circuit² has the property that the arbitrary initial state $|\psi\rangle$ of qubit A is "teleported" to the state in which qubit λ is left after the process. In the original teleportation scheme [20], two classical bits (resulting from a Bell measurement) are sent by the emitter, while the receiver performs a specific unitary operation on λ depending on these two bits. However, it is shown in [21] that these unitary operations can be performed at the quantum level as well, by using quantum logic gates and postponing the measurement of the two bits to the end of the circuit. The resulting quantum circuit (Fig. 4) is *formally* equivalent to the original teleportation scheme (although no classical bits are communicated) as exactly the same unitary transformations and quantum gates are involved. While we do not claim that an optical realization gives rise to "genuine" teleportation, this example circuit is instructive to demonstrate

² This teleportation circuit is equivalent to the one described in [19].

Fig. 4. Quantum circuit for teleportation (from [19]). The initial state of qubit Λ is teleported to the state of qubit λ . Qubits σ and λ must be initially in state $|0\rangle$. Qubits Λ and σ , if measured at the end of the circuit, yield two classical (random) bits that are uniformly distributed.



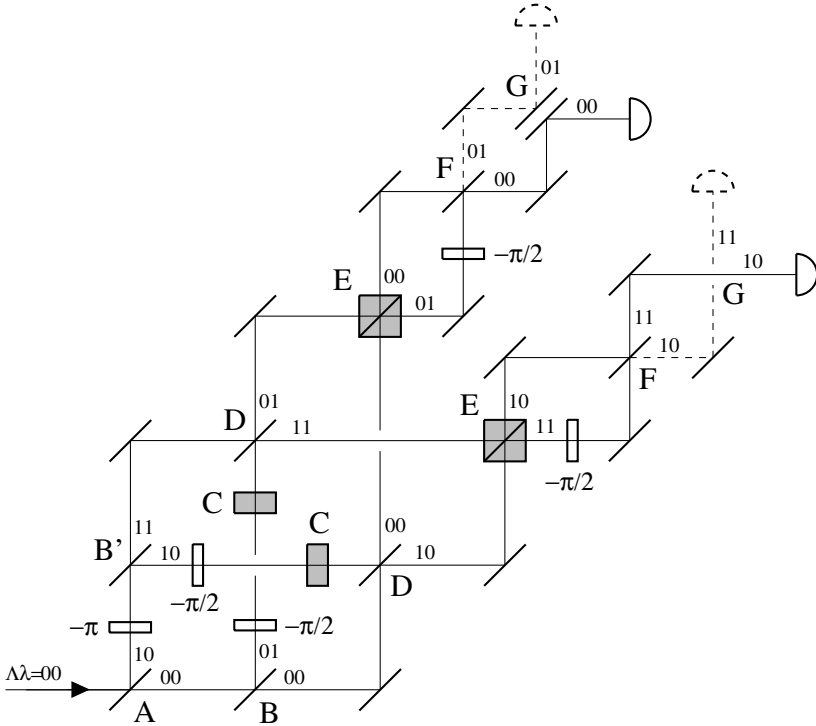
the correspondence between quantum logic and optical devices as it is small (3 qubits) but non-trivial.

In the optical counterpart of this circuit (see Fig. 5), qubits Λ and λ correspond to the location of the photon at the first and second splitting level, while σ stands for the polarization qubit. Note that the photons are initially horizontally polarized, *i.e.*, in polarization state $|0\rangle$. The first beam splitter A in Fig. 5 acts as a Hadamard gate on Λ , as explained previously. For convenience, we depict the teleportation of state $|\psi\rangle = |0\rangle$, so that the incident photon enters this beam splitter in the input port labeled $|0\rangle$. However, as any operation in $U(2)$ can be realized optically, an *arbitrary* state of Λ can be prepared (and then teleported) by having an additional beam splitter (with tunable phase shifters) connected to both input ports of beam splitter A. The second level of beam splitters B (and B')³ corresponds to the Hadamard gate B on λ in Fig. 4. The four paths at this point ($\Lambda\lambda = 00, 01, 10$, and 11) label the four components of the state vector characterizing qubits Λ and λ . The probability amplitude for observing the photon in each of these four paths, given the fact the photon enters the $|0\rangle$ port of beam splitters A and B, is then simply the corresponding component of the wave vector. The combined action of both controlled-NOT gates C in Fig. 4 is to flip the polarization state of the photon (qubit σ) conditionally on the parity of $\Lambda + \lambda \pmod{2}$, which is achieved by inserting polarization rotators C at the appropriate positions. In other words, the polarization is flipped on path 01 or 10, while it is unchanged on path 00 or 11.

The Hadamard gate D in Fig. 4 acts on qubit Λ , independently of λ . This is achieved in Fig. 5 by grouping the paths in pairs with the same value of λ (*i.e.*, crossing the paths) and using two beam splitters D in order to effect a Hadamard transformation on Λ (one for each value of λ). Similarly, the controlled-NOT gate E acting on λ (conditionally on the polarization) is implemented by the use

³ For convenience, two realizations (B and B') of the Hadamard gate are used in Fig. 5, where B' is obtained from B by interchanging the $|0'\rangle$ and $|1'\rangle$ output ports in Fig. 11a.

Fig. 5. Optical realization of the quantum circuit for teleportation using polarized photons. The location qubit λ characterizes the “which-arm” information at the first beam splitter, while qubit λ stands for the “which-path” information at the second level of splitting. The initial location qubit λ is teleported to qubit λ and probed via the interference pattern observed at the upper or lower ($\lambda = 0, 1$) final beam splitter, for both polarization states ($\sigma = 0, 1$) of the detected photon.



of two polarizing beam splitters E after crossing the paths again. A polarizing beam splitter leaves a horizontally polarized photon (state $|0\rangle$) unchanged, while vertical polarization (state $|1\rangle$) is reflected. The last Hadamard gate F in Fig. 4 corresponds to the last two beam splitters F, and the final controlled-NOT gate G is simply achieved by crossing the paths ($\lambda = 0, 1$) in the lower arm ($\lambda = 1$) versus the upper arm ($\lambda = 0$). In fact, the setup could be simplified by noting that the conditional crossing of paths achieved by G simply reduces to relabeling the output ports of beam splitter F in the $\lambda = 1$ arm. In Fig. 5, only those phase shifters associated with the Hadamard gates (Fig. 1a) that are relevant in the final detection are indicated.

The interpretation of this optical circuit in the language of teleportation is the following. After being “processed” in this quantum circuit, a photon which was initially horizontally polarized can reach one of the two “light” detectors

(solid line in Fig. 5) with horizontal or vertical polarization. This corresponds to the final measurement of qubits Λ and σ in Fig. 4 yielding two classical (random) bits: upper or lower arm, horizontal or vertical polarization. The third qubit, λ , contains the teleported quantum bit, that is, the initial arbitrary state of Λ . Since the location state of the photon is initially $|0\rangle$ in the setup represented in Fig. 5 it always exits to the “light” detector and never reaches the “dark” one (dashed line). For any measured value of Λ (photon detected in the upper or lower arm) and σ (horizontally or vertically polarized photon), the entire setup forms a simple balanced Mach-Zehnder interferometer. Indeed, there are exactly two *indistinguishable* paths leading to each of the eight possible outcomes (four detectors, two polarizations); these interfere pairwise, just as in a standard Mach-Zehnder interferometer, explaining the fact that the photon always reaches the “light” detector (in both $\Lambda = 0$ and $\Lambda = 1$ arms and for both polarizations). In this sense, the initial “which-arm” qubit Λ has been teleported to the final “which path” qubit λ . Note that, as no photodetection coincidence is required in this optical experiment, the setup is actually not limited to *single*-photon interferometry. This largely simplifies the realization of the optical source since classical light fields (such as those from a laser) can be used rather than number states.

4 Conclusion

We have proposed a general technique for simulating small-scale quantum networks using optical setups composed of linear optical elements only. This avoids the recourse to non-linear Kerr media to effect quantum conditional dynamics, a severe constraint in the usual optical realization of quantum circuits. A drawback of this technique is clearly the exponential increase of the resources (optical devices) with the size of the circuit. Nevertheless, as optical components that simulate 1- and 2-bit universal quantum gates can be cascaded straightforwardly, a non-trivial quantum computing optical device can easily be constructed if the number of component qubits is not too large. We believe this technique can be applied without fundamental difficulties to the encoding and decoding circuits that are involved in the simplest quantum error-correcting schemes [22], opening up the possibility for an experimental simulation of them. Furthermore, this technique promises a technologically simple way to test quantum algorithms for performance and error stability. Last but not least, the correspondence between quantum circuits and optical (interferometric) setups suggests that new and improved interferometers could be designed using the quantum circuit language [23].

Acknowledgments

We thank Paul Kwiat for many discussions and collaboration in this project. This work was supported in part by NSF Grants PHY 94-12818 and PHY 94-20470,

and by a grant from DARPA/ARO through the QUIC Program (#DAAH04-96-1-3086). N.J.C. is Collaborateur Scientifique of the Belgian National Fund for Scientific Research.

References

1. P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, in *Proc. of the 35th Annual Symposium on Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, New York, 1994), pp. 124-134.
2. D. G. Cory, M. D. Price, and T. F. Havel, Nuclear magnetic resonance spectroscopy: An experimentally accessible paradigm for quantum computing, eprint quant-ph/9709001 (to appear in *Physica D*); N. Gershenfeld and I. L. Chuang, Bulk spin-resonance quantum computation, *Science* **275** (1997) 350; I. L. Chuang, N. Gershenfeld, M. G. Kubinec, and D. W. Leung, Bulk quantum computation with nuclear-magnetic resonance—Theory and experiment. *Proc. Roy. Soc. London A* **454** (1998) 447.
3. C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, Demonstration of a fundamental quantum logic gate, *Phys. Rev. Lett.* **75** (1995) 4714; Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, and H. J. Kimble, Measurement of conditional phase shifts for quantum logic, *Phys. Rev. Lett.* **75** (1995) 4710.
4. K. Obenland and A. Despain, these proceedings.
5. C. Miquel, J. P. Paz, and W. H. Zurek, Quantum computation with phase drift errors, *Phys. Rev. Lett.* **78** (1997) 3971.
6. D. G. Cory et al., Experimental quantum error correction, eprint quant-ph/9802018.
7. I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd, Experimental realization of a quantum algorithm, *Nature* **393**, 143 (1998).
8. D. P. DiVincenzo, Quantum computation, *Science* **270**, 255 (1995).
9. A. Barenco *et al.*, Elementary gates for quantum computation, *Phys. Rev. A* **52**, 3457 (1995).
10. D. P. DiVincenzo, *Phys. Rev. A* **51**, 1015 (1995); D. Deutsch, A. Barenco, and A. Ekert, *Proc. R. Soc. London A* **449**, 669 (1995); S. Lloyd, *Phys. Rev. Lett.* **75**, 346 (1995).
11. M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental realization of any discrete unitary operator, *Phys. Rev. Lett.* **73**, 58 (1994).
12. B. Yurke, S. L. McCall, and J. R. Klauder, SU(2) and SU(1,1) interferometers, *Phys. Rev. A* **33**, 4033 (1986); S. Prasad, M. O. Scully, and W. Martienssen, A quantum description of the beam splitter, *Opt. Commun.* **62**, 139 (1987).
13. I. L. Chuang and Y. Yamamoto, Simple quantum computer, *Phys. Rev. A* **52**, 3489 (1995); Quantum bit regeneration, *Phys. Rev. Lett.* **76**, 4281 (1996).
14. R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics Vol. III* (Addison-Wesley, Reading, MA, 1965)
15. V. Degiorgio, Phase shift between the transmitted and the reflected optical fields of a semireflecting lossless mirror is $\pi/2$, *Am. J. Phys.* **48**, 81 (1980); A. Zeilinger, General properties of lossless beam splitters in interferometry, *ibid.* **49**, 882 (1981); Z. Y. Ou and L. Mandel, Derivation of reciprocity relations for a beam splitter from energy balance, *ibid.* **57**, 66 (1989).

16. N. J. Cerf, C. Adami, and P. G. Kwiat, Optical simulation of quantum logic, *Phys. Rev. A* **57**, R1477 (1998).
17. G. J. Milburn, Quantum optical Fredkin gate, *Phys. Rev. Lett.* **62**, 2124 (1989).
18. M. O. Scully and K. Drühl, Quantum eraser—A proposed photon-correlation experiment concerning observation and delayed choice in quantum mechanics. *Phys. Rev. A* **25**, 2208 (1982); B.-G. Englert, M. O. Scully, and H. Walther, The duality in matter and light, *Sci. Am.* **271**(6), 86 (1994).
19. G. Brassard, S. L. Braunstein, and R. Cleve, *Physica D* (1998), to appear.
20. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.* **70**, 1895 (1993).
21. S. L. Braunstein, Quantum teleportation without irreversible detection, *Phys. Rev. A* **53** (1996) 1900.
22. L. Vaidman, L. Goldenberg, and S. Wiesner, Error prevention scheme with four particles, *Phys. Rev. A* **54**, R1745 (1996); S. L. Braunstein, Quantum error correction of dephasing in 3 qubits, eprint quant-ph/9603024; R. Laflamme, C. Miquel, J. P. Paz, and W.H. Zurek, Perfect quantum error-correcting code, *Phys. Rev. Lett.* **77**, 198 (1996).
23. J. P. Dowling, Correlated input-port, matter-wave interferometer: Quantum-noise limits to the atom-laser gyroscope, *Phys. Rev. A* **57**, 4736 (1998).

Decoherence Control for Optical Qubits

D. Vitali and P. Tombesi

Dipartimento di Matematica e Fisica, Università di Camerino
via Madonna delle Carceri I-62032 Camerino, Italy

Abstract. Photons in cavities have been already used for the realization of simple quantum gates [Q.A. Turchette *et al.*, Phys. Rev. Lett. **75**, 4710 (1995)]. We present a method for combatting decoherence in this case.

1 Introduction

Quantum optics is usually concerned with the generation of nonclassical states of the electromagnetic field and their experimental detection. However with the recent rapid progress in the theory of quantum information processing the *protection* of quantum states and their quantum dynamics also is becoming a very important issue. In fact what makes quantum information processing much more attractive than its classical counterpart is its capability of using entangled states and of processing generic linear superpositions of input states. The entanglement between a pair of systems is capable of connecting two observers separated by a space-like interval, it can neither be copied nor eavesdropped on without disturbance, nor can it be used by itself to send a classical message [1]. The possibility of using linear superposition states has given rise to quantum computation, which is essentially equivalent to have massive parallel computation [2]. However all these applications crucially rely on the possibility of maintaining quantum coherence, that is, a defined phase relationship between the different components of linear superposition states, over long distances and for long times. This means that one has to minimize as much as possible the effects of the interaction of the quantum system with its environment and, in particular, decoherence, i.e., the rapid destruction of the phase relation between two quantum states of a system caused by the entanglement of these two states with two different states of the environment [3,4].

Quantum optics is a natural candidate for the experimental implementation of quantum information processing systems, thanks to the recent achievements in the manipulation of single atoms, ions and single cavity modes. In fact two quantum gates have been already demonstrated [5,6] in quantum optical systems and it would be very important to develop strategies capable of *controlling the decoherence* in experimental situations such as those described in Refs. [5,6].

In this paper we propose a simple physical way to control decoherence and protect a given quantum state against the destructive effects of the interaction with the environment: applying an appropriate feedback. The feedback scheme considered here has a quantum nature, since it is based on the injection of an

appropriately prepared atom in the cavity and some preliminary aspects of the scheme, and its performance, have been described in Refs. [7,8]

2 A Feedback Loop for Optical Cavities

Applying a feedback loop to a quantum system means subjecting it to a series of measurements and then using the result of these measurements to modify the dynamics of the system. Very often the system is continuously monitored and the associated feedback scheme provides a continuous control of the quantum dynamics. An example is the measurement of an optical field mode, such as photodetection and homodyne measurements, and for these cases, Wiseman and Milburn have developed a quantum theory of continuous feedback [9]. This theory has been applied in Refs. [10] to show that homodyne-mediated feedback can be used to slow down the decoherence of a Schrödinger cat state in an optical cavity.

Here we propose a different feedback scheme, based on direct photodetection rather than homodyne detection. The idea is very simple: whenever the cavity loses a photon, a feedback loop supplies the cavity mode with another photon, through the injection of an appropriately prepared atom. This kind of feedback is naturally suggested by the quantum trajectory picture of a decaying cavity field [11], in which time evolution is driven by the non-unitary evolution operator $\exp\{-\gamma ta^\dagger a/2\}$ interrupted at random times by an instantaneous jump describing the loss of a photon. The proposed feedback almost instantaneously “cures” the effect of a quantum jump and is able therefore to minimize the destructive effects of dissipation on the quantum state of the cavity mode.

In more general terms, the application of a feedback loop modifies the master equation of the system and therefore it is equivalent to an effective modification of the dissipative environment of the cavity field. For example, Ref. [12] shows that a squeezed bath [13] can be simulated by the application of a feedback loop based on a quantum non-demolition (QND) measurement of a quadrature of a cavity mode. In other words, feedback is the main tool for realizing, in the optical domain, the so called “quantum reservoir engineering” [14].

The master equation for continuous feedback has been derived by Wiseman and Milburn [9], and, in the case of perfect detection via a single loss source, is given by

$$\dot{\rho} = \gamma\Phi(a\rho a^\dagger) - \frac{\gamma}{2}a^\dagger a\rho - \frac{\gamma}{2}\rho a^\dagger a, \quad (1)$$

where γ is the cavity decay rate and $\Phi(\rho)$ is a generic superoperator describing the effect of the feedback atom on the cavity state ρ . Eq. (1) assumes perfect detection, i.e., all the photons leaving the cavity are absorbed by a unit-efficiency photodetector and trigger the cavity loop. It is practically impossible to realize such an ideal situation and therefore it is more realistic to generalize this feedback master equation to the situation where only a fraction $\eta < 1$ of the photons leaking out of the cavity is actually detected and switches on the atomic injector. It is immediate to see that (1) generalizes to

$$\dot{\rho} = \eta\gamma\Phi(a\rho a^\dagger) + (1 - \eta)\gamma a\rho a^\dagger - \frac{\gamma}{2}a^\dagger a\rho - \frac{\gamma}{2}\rho a^\dagger a. \quad (2)$$

Now, we have to determine the action of the feedback atom on the cavity field $\Phi(\rho)$; this atom has to release exactly one photon in the cavity, possibly regardless of the field state in the cavity. In the optical domain this could be realized using *adiabatic transfer of Zeeman coherence* [15].

2.1 Adiabatic Passage in a Three Level Λ Atom

A scheme based on the adiabatic passage of an atom with Zeeman substructure through overlapping cavity and laser fields has been proposed [15] for the generation of linear superpositions of Fock states in optical cavities. This technique allows for coherent superpositions of atomic ground state Zeeman sublevels to be “mapped” directly onto coherent superpositions of cavity-mode number states. If one applies this scheme in the simplest case of a three-level Λ atom one obtains just the feedback superoperator we are looking for, that is

$$\Phi(\rho) = a^\dagger(aa^\dagger)^{-1/2}\rho(aa^\dagger)^{-1/2}a, \quad (3)$$

corresponding to the feedback atom releasing exactly one photon into the cavity, regardless the state of the field.

To see this, let us consider a three level Λ atom with two ground states $|g_1\rangle$ and $|g_2\rangle$, coupled to the excited state $|e\rangle$ via, respectively, a classical laser field $\Omega(t)$ of frequency ω_L , and a cavity field mode of frequency ω . The corresponding Hamiltonian is

$$\begin{aligned} H(t) = & \hbar\omega a^\dagger a + \hbar\omega_{eg}|e\rangle\langle e| - i\hbar g(t)(|e\rangle\langle g_2|a - |g_2\rangle\langle e|a^\dagger) \\ & + i\hbar\Omega(t)(|e\rangle\langle g_1|e^{-i\omega_L t} - |g_1\rangle\langle e|e^{i\omega_L t}). \end{aligned} \quad (4)$$

The time dependence of $\Omega(t)$ and $g(t)$ is provided by the motion of the atom across the laser and cavity profiles. This Hamiltonian couples only states within the three-dimensional manifold spanned by $|g_1, n\rangle$, $|e, n\rangle$, $|g_2, n+1\rangle$, where n denotes a Fock state of the cavity mode. Of particular interest within this manifold is the eigenstate corresponding to the adiabatic energy eigenvalue (in the frame rotating at the frequency ω) $E_n = n\hbar\omega$,

$$|E_n(t)\rangle = \frac{g(t)\sqrt{n+1}|g_1, n\rangle + \Omega(t)|g_2, n+1\rangle}{\sqrt{\Omega^2(t) + (n+1)g^2(t)}} \quad (5)$$

which does not contain any contribution from the excited state and for this reason is called the “dark state”. This eigenstate exhibits the following asymptotic behavior as a function of time

$$|E_n\rangle \rightarrow \begin{cases} |g_1, n\rangle & \text{for } \Omega(t)/g(t) \rightarrow 0 \\ |g_2, n+1\rangle & \text{for } g(t)/\Omega(t) \rightarrow 0 \end{cases} \quad (6)$$

Now, according to the adiabatic theorem [16], when the evolution from time t_0 to time t_1 is sufficiently slow, a system starting from an eigenstate of $H(t_0)$ will

pass into the corresponding eigenstate of $H(t_1)$ that derives from it by continuity. This means that if the atom crossing is such that adiabaticity is satisfied, when the atom enters the interaction region in the ground state $|g_1\rangle$, the following adiabatic transformation of the atom-cavity system state takes place

$$\begin{aligned} & |g_1\rangle\langle g_1| \otimes \sum_{n,m} \rho_{n,m} |n\rangle\langle m| \\ & \rightarrow |g_2\rangle\langle g_2| \otimes \sum_{n,m} \rho_{n,m} |n+1\rangle\langle m+1| \\ & = |g_2\rangle\langle g_2| \otimes a^\dagger (aa^\dagger)^{-1/2} \rho (aa^\dagger)^{-1/2} a . \end{aligned} \quad (7)$$

Roughly speaking, this transformation amounts to a single photon transfer from the classical laser field to the quantized cavity mode realized by the crossing atom, provided that a counterintuitive pulse sequence in which the classical laser field $\Omega(t)$ is time-delayed with respect to $g(t)$ is applied. Figure 1 shows a simple diagram of the feedback scheme, together with the appropriate atomic configuration-

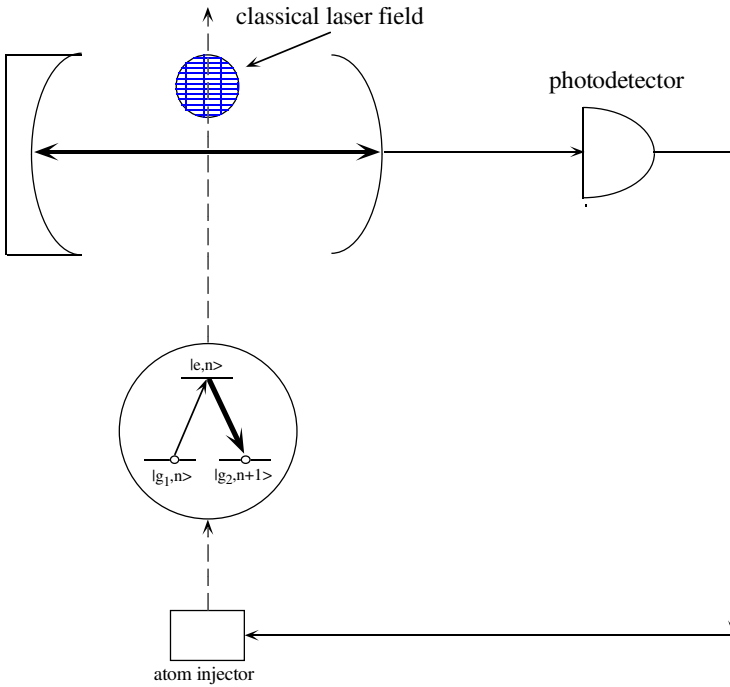


Fig. 1. Schematic diagram of the photodetection-mediated feedback scheme proposed for optical cavities, together with the appropriate atomic configuration for the adiabatic transfer.

uration, cavity and laser field profiles needed for the adiabatic transformation considered.

The quantitative conditions under which adiabaticity is satisfied are obtained from the requirement that the transition from the dark state $|E_n(t)\rangle$ to the other states be very small. One obtains [15,17]

$$\Omega_{max}, g_{max} \gg T_{cross}^{-1}, \quad (8)$$

where T_{cross} is the cavity crossing time and Ω_{max}, g_{max} are the two peak intensities.

The above arguments completely neglect dissipative effects due to cavity losses and atomic spontaneous emission. For example, cavity dissipation couples a given manifold $|g_1, n\rangle, |e, n\rangle, |g_2, n+1\rangle$ with those with a smaller number of photons. Since ideal adiabatic transfer occurs when the passage involves a single manifold, optimization is obtained when the photon leakage through the cavity is negligible during the atomic crossing, that is

$$T_{cross}^{-1} \gg \bar{n}\gamma, \quad (9)$$

where \bar{n} is mean number of photons in the cavity. On the contrary, the technique of adiabatic passage is robust against the effects of spontaneous emission as, in principle, the excited atomic state $|e\rangle$ is never populated. Of course, in practice some fraction of the population does reach the excited state and hence large values of g_{max} and Ω_{max} relative to the spontaneous emission rate γ_e are desirable. To summarize, the quantitative conditions for a practical realization of the adiabatic transformation (7) are

$$\Omega_{max}, g_{max} \gg T_{cross}^{-1} \gg \bar{n}\gamma, \gamma_e, \quad (10)$$

which, as pointed out in [15], could be realized in optical cavity QED experiments.

We note that when the adiabaticity conditions (10) are satisfied, then also the Markovian assumptions at the basis of the feedback master equation (2) are automatically justified. In fact, the continuous feedback theory of Ref. [9] is a Markovian theory derived assuming that the delay time associated to the feedback loop can be neglected with respect to the typical timescale of the cavity mode dynamics. In the present scheme the feedback delay time is due to the electronic trasmission time of the detection signal and, most importantly, by the interaction time T_{cross} of the atoms with the field, while the typical timescale of the cavity field dynamics is $1/\gamma\bar{n}$. Therefore, the inequality on the right of Eq. (10) is essentially the condition for the validity of the Markovian approximation and this *a posteriori* justifies our use of the Markovian feedback master equation (2) from the beginning.

2.2 Properties of the Adiabatic Transfer Feedback Model

When we insert the explicit expression (3) of the feedback superoperator into Eq. (2), the feedback master equation can be rewritten in the more transparent form

$$\dot{\rho} = \frac{(1-\eta)\gamma}{2} (2a\rho a^\dagger - a^\dagger a \rho - \rho a^\dagger a) - \frac{\eta\gamma}{2} [\sqrt{\hat{n}}, [\sqrt{\hat{n}}, \rho]] \quad (11)$$

that is, a standard vacuum bath master equation with effective damping coefficient $(1-\eta)\gamma$ plus an unconventional phase diffusion term, in which the photon number operator is replaced by its square root and which can be called “square root of phase diffusion”.

In the ideal case $\eta = 1$, vacuum damping vanishes and only the unconventional phase diffusion survives. As shown in Ref. [18], this is equivalent to say that ideal photodetection feedback is able to transform standard photodetection into a quantum non-demolition (QND) measurement of the photon number. In this ideal case, a generic Fock state $|n\rangle$ is obviously preserved for an infinite time, since each photon lost by the cavity triggers the feedback loop which, in a negligible time, is able to give the photon back through adiabatic transfer. However, the photon injected by feedback has no phase relationship with the photons already present in the cavity and, as shown by (11), this results in phase diffusion. This means that feedback does not guarantee perfect state protection for a generic *superposition of number states*, even in the ideal condition $\eta = 1$. In fact in this case, only the diagonal matrix elements in the Fock basis of the initial pure state are perfectly conserved, while the off-diagonal ones always decay to zero, ultimately leading to a phase-invariant state. However this does not mean that the proposed feedback scheme is good for preserving number states only, because the unconventional “square-root of phase diffusion” is much slower than the conventional one (described by a double commutator with the number operator).

In fact the time evolution of a generic density matrix element in the case of feedback with ideal photodetection $\eta = 1$ is

$$\rho_{n,m}(t) = \exp \left\{ -\frac{\gamma t}{2} (\sqrt{n} - \sqrt{m})^2 \right\} \rho_{n,m}(0) , \quad (12)$$

while the corresponding evolution in the presence of standard phase diffusion is

$$\rho_{n,m}(t) = \exp \left\{ -\frac{\gamma t}{2} (n - m)^2 \right\} \rho_{n,m}(0) . \quad (13)$$

Since

$$(n - m)^2 \geq (\sqrt{n} - \sqrt{m})^2 = \frac{(n - m)^2}{(\sqrt{n} + \sqrt{m})^2} \quad \forall n, m \quad (14)$$

each off-diagonal matrix element decays slower in the square root case and this means that the feedback-induced unconventional phase diffusion is slower than the conventional one.

A semiclassical estimation of the diffusion constant can be obtained from the time evolution of the mean coherent amplitude $\langle a(t) \rangle$. In fact, phase diffusion causes a decay of this amplitude as the phase spreads around 2π , even if the photon number is conserved. In the presence of ordinary phase diffusion the amplitude decays at the rate $\gamma/2$; in fact

$$\langle a(t) \rangle = \text{Tr} \{ a \rho(t) \} = \sum_{n=0}^{\infty} \sqrt{n+1} \rho_{n+1,n}(t) , \quad (15)$$

and using Eq. (13) one gets

$$\langle a(t) \rangle = e^{-\gamma t/2} \langle a(0) \rangle .$$

In the case of the square root of phase diffusion, Eqs. (12) and (15) instead yield

$$\langle a(t) \rangle = \text{Tr} \{ a(t) \rho(0) \} , \quad (16)$$

where the Heisenberg-like time evolved amplitude operator $a(t)$ is given by

$$a(t) = \exp \left\{ -\frac{\gamma t}{2} \left(\sqrt{a a^\dagger} - \sqrt{a^\dagger a} \right)^2 \right\} a . \quad (17)$$

In the semiclassical limit it is reasonable to assume a complete factorization of averages, so to get

$$\langle a(t) \rangle = \exp \left\{ -\frac{\gamma t}{2} \left(\sqrt{\bar{n}+1} - \sqrt{\bar{n}} \right)^2 \right\} \langle a(0) \rangle , \quad (18)$$

which, in the limit of large mean photon number \bar{n} , yields

$$\langle a(t) \rangle = \exp \left\{ -\frac{\gamma t}{8\bar{n}} \right\} \langle a(0) \rangle . \quad (19)$$

This slowing down of phase diffusion (similar to that taking place in a laser well above threshold) means that, when the feedback efficiency η is not too low, the “lifetime” of generic pure quantum states of the cavity field can be significantly increased with respect to the standard case with no feedback (see Eq. (11)).

3 Optical Feedback Scheme for the Protection of Qubits

Photon states are known to retain their phase coherence over considerable distances and for long times and for this reason high-Q optical cavities have been proposed as a promising example for the realization of simple quantum circuits for quantum information processing. To act as an information carrying quantum state, the electromagnetic fields must consist of a superposition of few distinguishable states. The most straightforward choice is to consider the superposition of the vacuum and the one photon state $\alpha|0\rangle + \beta|1\rangle$. However it is easy to understand that this is not convenient because any interaction coupling $|0\rangle$ and $|1\rangle$ also couples $|1\rangle$ with states with more photons and this leads to information losses. Moreover the vacuum state is not easy to observe because it cannot be distinguished from a failed detection of the one photon state. A more convenient and natural choice is *polarization coding*, i.e., using two degenerate polarized modes and qubits of the following form

$$|\psi\rangle = \left(\alpha a_+^\dagger + \beta a_-^\dagger \right) |0\rangle = \alpha |0, 1\rangle + \beta |1, 0\rangle , \quad (20)$$

in which one photon is shared by the two modes [21]. In fact this is a “natural” two-state system, in which the two basis states can be easily distinguished with polarization measurements; moreover they can be easily transformed into each other using polarizers.

Polarization coding has been already employed in one of the few experimental realization of a quantum gate, the quantum phase gate realized at Caltech [5]. This experiment has demonstrated conditional quantum dynamics between two frequency-distinct fields in a high-finesse optical cavity. The implementation of this gate employs two single-photon pulses with frequency separation large compared to the individual bandwidth, and whose internal state is specified by the circular polarization basis as in (20). The conditional dynamics between the two fields is obtained through an effective strong Kerr-type nonlinearity provided by a beam of cesium atoms. This conditional dynamics of the quantum gate has to be unitary with a high degree of accuracy during the operation time, i.e., decoherence has to be negligible; the experiment of Ref. [5] has been performed in the bad cavity limit, in which the main dissipative effects and main source of decoherence is the photon leakage outside the cavity. It is therefore quite natural to see if the atomic feedback scheme described in detail above is able to protect the “flying” qubits of Ref. [5]. To be more specific, we shall not be concerned with the protection of the quantum gate dynamics, but we shall focus on a simpler but still important problem: protecting an unknown input state for the longest possible time against decoherence. We shall therefore consider a single qubit, i.e., a single frequency whose internal state is specified by the polarization.

One has to apply an adiabatic transfer feedback loop as that of Fig. 1 to each polarized mode independently. This can be done using polarization-sensitive detectors (for example a polarized beam splitter and two detectors) and two classical laser fields with opposite circular polarization. In this way one has two similar feedback loops where one polarized mode is involved in the transition $|g_1\rangle \rightarrow |g_2\rangle$, and the other mode participates to the reversed transition. In this way each mode gets a photon with the right polarization. A schematic description of the scheme is given by Fig. 2.

For a quantitative characterization of how the feedback scheme is able to protect an initial pure state we study the fidelity $F(t)$

$$F(t) = \text{Tr} \{ \rho(0) \rho(t) \} \quad (21)$$

i.e., the overlap between the final and the initial state $\rho(0)$ after a time t . In general $0 \leq F(t) \leq 1$. For an initially pure state $|\psi(0)\rangle$, $F(t)$ is in fact the probability to find the system in the initial state at a later time. A decay to an asymptotic limit is given by the overlap $\langle \psi(0) | \rho(\infty) | \psi(0) \rangle$. Since the input state we seek to protect is unknown, the protection capabilities of the feedback scheme are better characterized by the minimum fidelity, i.e., the fidelity of Eq. (21) minimized over all possible initial states. Moreover we shall consider a class of initial states more general than those of Eq. (20), i.e.,

$$|\psi\rangle = \alpha|n, m\rangle + \beta|m, n\rangle. \quad (22)$$

Since the two polarized modes evolve independently, one has to solve the master equation (11) to calculate the fidelity. This can be done easily and one gets the following expression for the minimum fidelity

$$F_{min}(t) = \frac{1}{2} \left(e^{-(1-\eta)\gamma t(n+m)} + e^{-\gamma t(n+m-2\eta\sqrt{nm})} \right). \quad (23)$$

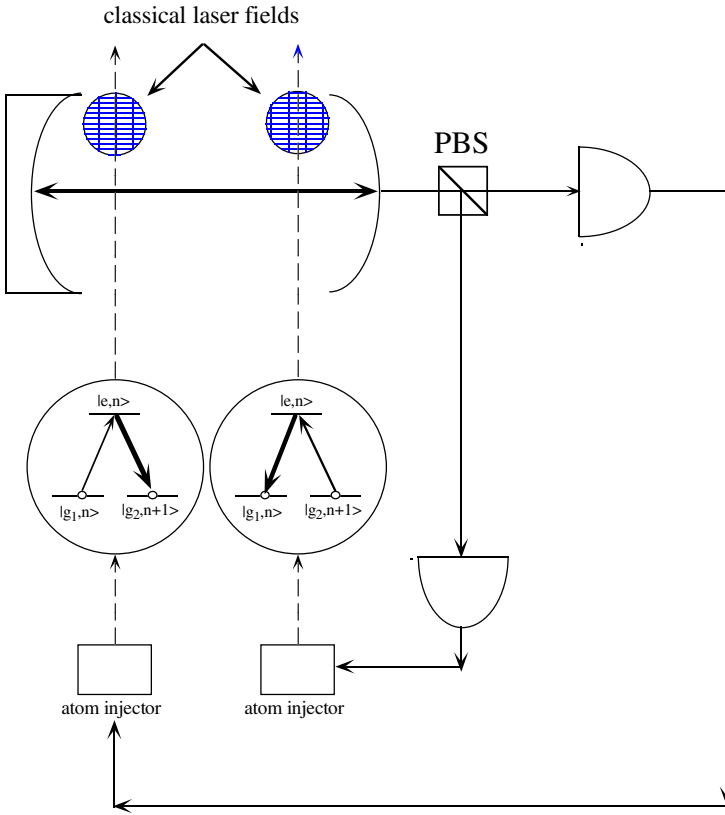


Fig. 2. Adaptation of the feedback scheme of Fig. 1 to the polarization coding case; there is a feedback loop for each circularly polarized mode and the two loops are separated by the polarized beam splitter PBS.

In the absence of feedback ($\eta = 0$), this expression becomes $F_{min}(t) = \exp\{-\gamma t(n+m)\}$ showing that in this case, the states most robust against cavity damping are those with the smallest number of photons, $m+n=1$, i.e., the states of the form of Eq. (20). Moreover, in a typical quantum information processing situation, one has to consider small qubit “storage” times t with respect to γ^{-1} so to have reasonably small error probabilities in quantum information

storage. Therefore the protection capability of an optical cavity with no feedback applied is described by

$$F_{min}(t) = 1 - \gamma t. \quad (24)$$

If we now consider the situation in the presence of feedback (Eq. (23)), it is possible to see that, for fixed, non-unit efficiency η , the best protected state are, as in the no-feedback case, the states with only one photon $\alpha|0, 1\rangle + \beta|1, 0\rangle$ and therefore the corresponding minimum fidelity for $\eta < 1$ is given by

$$F_{min}(t) = \frac{1}{2} \left(e^{-(1-\eta)\gamma t} + e^{-\gamma t} \right) \simeq 1 - \gamma t \left(1 - \frac{\eta}{2} \right). \quad (25)$$

This shows that feedback increases the “lifetime” of a generic qubit state with respect to the no-feedback case, even if, in this non-ideal case, one has a scaling of the error probability by a factor $(1 - \eta/2)$ only.

It is interesting to consider the ideal case $\eta = 1$, even if it is not realistic, because in this case the situation changes qualitatively. In fact Eq. (23) becomes

$$F_{min}(t) = \frac{1}{2} \left(1 + e^{-\gamma t (\sqrt{n} - \sqrt{m})^2} \right). \quad (26)$$

so that it is easy to see that in this case it becomes convenient to work with a large number of photons and that the best protected qubit states are those of the form

$$|\psi\rangle = \alpha|n, n+1\rangle + \beta|n+1, n\rangle \quad n \gg 1 \quad (27)$$

whose corresponding minimum fidelity is

$$F_{min}(t) \simeq \frac{1}{2} \left(1 + e^{-\gamma t/4n} \right) \simeq 1 - \frac{\gamma t}{8n}.$$

Therefore, in the ideal photodetection case and using qubits of the form of (27), the probability of errors in the storage of quantum information can be made arbitrarily small.

The feedback method proposed here to deal with decoherence in quantum information processing is different from most of the proposals made in this research field, which are based on the so called quantum error correction codes [22]. In our case, feedback allows a physical control of decoherence, through a continuous monitoring and eventual correction of the dynamics and in this sense our approach is similar in spirit to the approach of Ref. [23,24]. Quantum error correction is instead a way to use *software* to preserve linear superposition states. Essentially in these approaches the entangled superposition state of l qubits is “encoded” in a larger number of qubits n , so that, assuming that only a fraction of qubits t/n decoheres, it is possible to reconstruct the original state with a suitable decoding procedure. However, due to existence of a lower (quantum Hamming) and an upper (quantum Gilbert-Varshamov) bound for the “code rate” l/n [25], these quantum error correction codes can be applied and become efficient only at sufficiently small probability of error t/n . For this reason, even if under realistic conditions our feedback scheme achieves only a moderate reduction of the error probability, it could be useful when used in *conjunction* with

quantum error correction techniques. The feedback scheme would realize the preliminary reduction of the error probability, which is necessary for an optimal implementation of efficient error correction schemes.

4 Acknowledgments

This work has been partially supported by the Istituto Nazionale Fisica della Materia (INFM) through the “Progetto di Ricerca Avanzata INFM-CAT”.

References

1. C.H. Bennett, in *Quantum Communication, Computing and Measurement*, edited by O. Hirota, A.S. Holevo and C.M. Caves (Plenum Press, New York, 1997), pag. 25.
2. A. Ekert and R. Josza, *Rev. Mod. Phys.* **68**, 733 (1996).
3. W.H. Zurek, *Phys. Today* **44**(10), 36 (1991), and references therein.
4. A.J. Leggett, in *Chance and Matter* (Proceedings, 1986 Les Houches Summer School), ed. by J. Souletie, J. Vannimenus, and R. Stora, (North Holland, Amsterdam, 1987), pag. 395.
5. Q.A. Turchette, C.J. Hood, W. Lange, H. Mabuchi and H.J. Kimble, *Phys. Rev. Lett.* **75**, 4710 (1995).
6. C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano and D.J. Wineland, *Phys. Rev. Lett.* **75**, 4714 (1995).
7. D. Vitali, P. Tombesi, G.J. Milburn, *Phys. Rev. Lett.* **79**, 2442 (1997).
8. D. Vitali, P. Tombesi, G.J. Milburn, *J. Mod. Opt.* **44**, 2033 (1997).
9. H.M. Wiseman and G.J. Milburn, *Phys. Rev. Lett.* **70**, 548 (1993); *Phys. Rev. A* **49**, 1350 (1994); H.M. Wiseman, *Phys. Rev. A* **49**, 2133 (1994).
10. P. Tombesi and D. Vitali, *Phys. Rev. A* **51**, 4913 (1995); P. Goetsch, P. Tombesi and D. Vitali, *Phys. Rev. A* **54**, 4519 (1996).
11. H.J. Carmichael, *An Open Systems Approach to Quantum Optics*, (Springer, Berlin, 1993).
12. P. Tombesi and D. Vitali, *Phys. Rev. A* **50**, 4253 (1994).
13. C.W. Gardiner, *Quantum Noise*, (Springer, Berlin, 1991).
14. J.F. Poyatos, J.I. Cirac and P. Zoller, *Phys. Rev. Lett.* **77**, 4728 (1996).
15. A.S. Parkins, P. Marte, P. Zoller, O. Carnal and H.J. Kimble, *Phys. Rev. A* **51**, 1578 (1995) and references therein.
16. A. Messiah, *Quantum Mechanics* (North Holland, Amsterdam, 1962).
17. J.R. Kuklinski, U. Gaubatz, F.T. Hioe and K. Bergmann, *Phys. Rev. A* **40**, 6741 (1990).
18. H.M. Wiseman, *Phys. Rev. A* **51**, 2459 (1995).
19. K. Vogel, V.M. Akulin and W.P. Schleich, *Phys. Rev. Lett.* **71**, 1816 (1993).
20. C.K. Law and J.H. Eberly, *Phys. Rev. Lett.* **76**, 1055 (1996).
21. S. Stenholm, *Opt. Comm.* **123**, 287 (1996); P. Törmä and S. Stenholm, *Phys. Rev. A* **54**, 4701 (1996).
22. E. Knill and R. Laflamme, *Phys. Rev. A* **55**, 900 (1997) and references therein.
23. T. Pellizzari, S.A. Gardiner, J.I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **75**, 3788 (1995).
24. H. Mabuchi and P. Zoller, *Phys. Rev. Lett.* **76**, 3108 (1996).
25. A. Ekert and C. Macchiavello, *Phys. Rev. Lett.* **77**, 2585 (1996).

Adiabatic Controlled-NOT Gate for Quantum Computation

D.V. Averin

Department of Physics and Astronomy, SUNY at Stony Brook,
Stony Brook, NY 11794

Abstract. A new variant of the controlled-NOT quantum logic gate is proposed. The gate is based on adiabatic level-crossing dynamics of the q-bits. An important advantage of the adiabatic dynamics, as opposed to the ac-driven Rabi transitions, is its considerable insensitivity to the unavoidable spread of the gate parameters. The gate has a natural implementation in terms of the Cooper pair transport in arrays of small Josephson tunnel junctions. The decoherence rate for this implementation is estimated.

1 Introduction

The invention of quantum algorithms and quantum information processing (see, e.g., the reviews [1]) changed the foundations of the theoretical computer science by demonstrating that the information is processed differently by quantum and classical systems. In an ideal world, a quantum algorithm implemented on a quantum computer can radically outperform the classical algorithm by making use of quantum parallelism inherent in entangled quantum states. Examples of problems which can be efficiently solved with quantum algorithms include factorization of large numbers [2] and database search [3].

Practical realization of a quantum computer requires, however, very precise and reversible time evolution of complex quantum mechanical systems, the fact that gives rise to serious doubts [4] as to whether even the simplest version of a quantum computer will ever become a reality. It is therefore important to look into various possible ways of implementing simple elements of quantum computer – quantum logic gates in order to find the optimal approach to building such a computer. Generally speaking, a quantum gate should satisfy two contradictory requirements: being isolated from the outside world in order to maintain quantum coherence, and interacting with other q-bits, read-out system, etc., in order to perform a meaningful computation. Existing quantum gate proposals use various systems including ion traps [5,6], electrodynamic cavities [7], semiconductor quantum dots [8,9], NMR spectroscopy [10], quantum flux dynamics in SQUIDS [11]. Some of these proposals, for instance, ion-trap or NMR, are characterized by potentially very long relaxation times, since the gates in these proposals are well isolated physically from the outside world. However, due to the very same reason, these gates can not be combined easily to form larger computing systems. For other gates, for instance, based on semiconductor quantum

dots, the situation is the opposite. In principle, it is not too difficult to integrate them into larger structures, but there seems to be very little hope of reducing decoherence rates to a level acceptable for quantum computation.

The aim of this work is to suggest another possible realization of quantum gates which is based on manipulation of the charge states of small Josephson tunnel junctions. This approach combines both the potential for relatively long relaxation times and large degree of design flexibility, and probably represents one of the best, if not the best, hope for realization of a quantum computer of medium complexity. Such a computer, while not being sufficient for factorization of large numbers of practical interest, could be sufficiently complex to perform privacy amplification [12,13] in quantum communication channels.

The basic universal set of quantum logic gates consists of the one q-bit gates and two q-bit controlled-NOT (CN) gate. In practically any implementation, including the one discussed below, the dynamics of the two q-bit gates contains all elements of the one q-bit operation, and therefore, we can concentrate on the two q-bit CN gate. The operation of this gate can be described simply as inversion of the target q-bit states when the control q-bit is the “1” state. The state of the control q-bit should be unchanged during this operation. Typically, the CN-operation is achieved through the use of the ac-driven Rabi transition between the q-bit states [5,6,10]. This approach becomes very problematic for solid state implementations of the quantum gates because of the unavoidable variations of parameters from gate to gate in solid state structures. We propose another general scheme of the CN gate which uses adiabatic transitions between the q-bit states and is more suitable for implementation in systems of small tunnel Josephson junctions. Although adiabatic approach does not solve completely the problem of parameter variations it makes it less severe.

2 Adiabatic CN-Gate

The main idea of the adiabatic CN-gate is as follows. Interaction between the control and target q-bit makes the energy difference between the two states of the target q-bit dependent on the state of the control q-bit. If the control q-bit is in the state “1” of the computational basis, the energy difference is smaller and under application of the time-dependent bias the target q-bit passes through the level-crossing point, where the energies of its two states are equal – see Fig. 1. If the rate of the bias increase is sufficiently small, the two states of the target q-bit exchange their occupation probabilities. When the control q-bit is in the “0” state, the energy difference is larger and the same bias pulse does not drive the target q-bit through the level-crossing point. In this case, the occupation probabilities remain the same. The tunnel coupling between the states of the control q-bit is suppressed during the entire process so that their occupation probabilities do not change in either case. This time evolution realizes CN-gate operation provided that the parameters of the two q-bits are chosen in such a way that the dynamic phases accumulated in the system evolution along all four “paths” are equal.

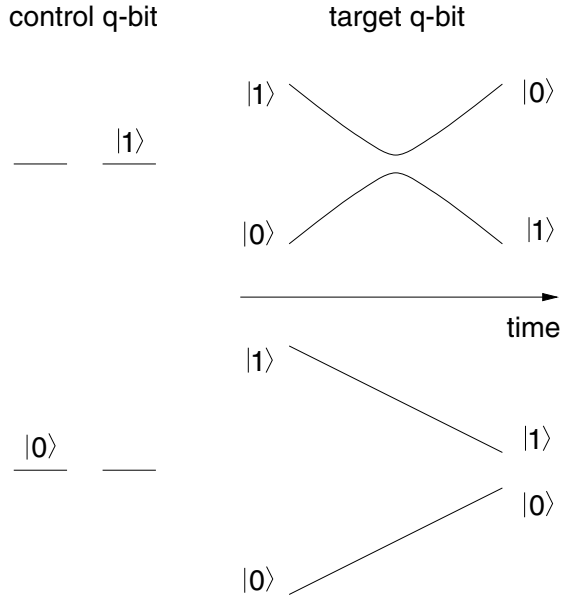


Fig. 1. Time evolution of the energy levels of a controlled-NOT quantum logic gate based on the adiabatic level-crossing dynamics.

To implement this dynamics of the two q-bit system we need to control both the energy difference $\varepsilon_j(t)$ between the two states of a q-bit in the computational basis $|0\rangle$ and $|1\rangle$, and the tunneling amplitude $\Omega_j(t)$ in this basis, i.e., the Hamiltonian of the system should be:

$$H = \sum_{j=1,2} (\varepsilon_j(t)\sigma_{zj} + \Omega_j(t)\sigma_{xj}) + \eta\sigma_{z1}\sigma_{z2}, \quad (1)$$

where σ_j are the Pauli matrices for the j th q-bit, and η is the energy of interaction between the q-bits. Although the basic gate dynamics does not require the modulation of the interaction strength η , this interaction does not allow to make the energies of all four gate states equal after the gate operation. This means that the relative phases of these states will continue to evolve at a rate on the order of η/\hbar and we need to be able to manipulate the gate on a time scale much shorter than \hbar/η , which in its turn should be much shorter than the time scale of the adiabatic dynamics. This long hierarchy of time scales presents a serious problem that exists for other proposals of quantum gates as well. A better solution is to design a gate in a way that allows to switch the interaction on and off, despite the fact that this makes the design appreciably more complicated.

If the interaction energy η in the Hamiltonian (1) can be controlled, we can separate the gate dynamics into three steps. At first, the two q-bits are brought into contact by switching on η and Ω_2 (Ω_1 is completely suppressed throughout the gate operation). Simultaneously the energy difference ε_2 between the states

of the target q-bit is set to some nonvanishing initial value. Then this energy difference is increased while all other energies are kept constant. This step is the central “level-crossing” part of the gate dynamics. During the final third step both η and Ω_2 are suppressed back to zero so that the two q-bits are effectively separated and ε_2 can also be reduced to zero.

The precise functional dependence of Ω_2 , ε_2 , and η on time does not qualitatively affect the gate dynamics, as long as all these parameters are changed gradually. The limitation on the rate of the parameter variations is associated with the unwanted transitions between the instantaneous energy eigenstates of the system which are brought about by these variations. These transitions violate the correct adiabatic dynamics which assumes that the system remains at all times in the same eigenstate it occupied initially. Adopting a simple model time dependence of the energy difference $\varepsilon_2(t)$:

$$\varepsilon_2(t) = \varepsilon + u \tanh(t/\tau), \quad (2)$$

and using the standard quasiclassical approach we can calculate explicitly the probability p that the system makes an unwanted transition during the central second step of the gate operation. This simple calculation confirms the expected result that the probability p reaches maximum when the system passes through the level crossing-point and is given then by the standard Landau-Zener expression:

$$p_{LZ} = \exp\left\{-\frac{\pi\tau\Omega^2}{\hbar u}\right\}. \quad (3)$$

Here Ω is the magnitude of the tunnel amplitude $\Omega_2(t)$ that is kept constant during this step of the gate operation. Thus, the condition $p_{LZ} \ll 1$, i.e., $\tau \gg \hbar u/\Omega^2$, ensures the correct adiabatic dynamics of the gate.

This implies that dynamics of the occupation probabilities of the states of adiabatic CN gate is not sensitive to the precise values of the parameters in the Hamiltonian (II) provided that they satisfy several constraints which ensure the gate operation shown in Fig. 1:

$$\eta + u - \varepsilon \gg \Omega, \quad \eta - u - \varepsilon \ll -\Omega, \quad u - \eta - \varepsilon \ll -\Omega. \quad (4)$$

If all these conditions are satisfied, the evolution of the absolute values of the occupation amplitudes α_{ij} of the four gate states corresponds to the correct CN operation:

$$|\alpha_{0j}| \rightarrow |\alpha_{0j}|, \quad |\alpha_{10}| \rightarrow |\alpha_{11}|, \quad |\alpha_{11}| \rightarrow |\alpha_{10}|.$$

(The indices i and j denote the states of the control and target q-bit respectively.) Besides this time evolution of the absolute values of α_{ij} , the correct gate dynamics requires also that phases of the four states accumulated in the process of the gate evolution are equal modulo 2π . This can be achieved by adjusting the bias $\varepsilon_{1,2}$ of the two q-bits and the energy splitting Ω_2 during the gate operation. The bias ε_1 controls the relative phases of the pairs of states evolving from the 0 and 1 state of the control q-bit, while ε_2 and Ω_2 control the phases within each pair. With such an adjustment of the phases, the adiabatic time evolution of the coupled q-bits represents correctly the CN quantum logic gate.

3 Small Josephson Junctions as Q-Bits

The adiabatic CN-gate can be naturally implemented in systems of small Josephson tunnel junctions in the Coulomb blockade regime – see, e.g., [14,15]. The energy diagram of an elementary building block of such a system, a single junction, is shown in Fig. 2. The dominant contribution to the junction energy is given by the charging energy $U(n)$ of the junction as a capacitor:

$$U(n) = \frac{(2en - Q_0)^2}{2C},$$

where C is the junction capacitance, n is the number of Cooper pairs transferred across the junction, and Q_0 is the charge induced by the external bias voltage V_0 across the junction, $Q_0 = V_0 C$. In general, the states with different n 's are separated by large energy gaps on the order of elementary charging energy $E_C = e^2/2C$. However, when the external voltage V_0 induces the charge of approximately one electron on the junction capacitance, $Q_0 \simeq e$, the two state, $n = 0$ and $n = 1$ are nearly degenerate and are separated from all other states by the large energy gaps – see Fig. 2. In this regime the junction behaves effectively as a two-level system. The energy difference ε between the level of this two-level system is controlled by the external voltage $\varepsilon = 2e(e - Q_0)/C$, while the amplitude Ω of tunneling between them is determined by the Josephson coupling energy E_J of the junction, $\Omega = E_J/2$. The Josephson coupling energy depends on the tunnel resistance R_T of the insulator barrier between the electrodes, and for the electrodes with equal superconducting energy gaps Δ is equal to $\pi\hbar\Delta/4e^2R_T$ – see, e.g., [15].

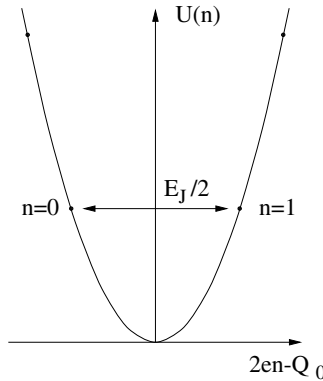


Fig. 2. Energy diagram of a tunnel Josephson junction in the Coulomb blockade regime biased with the external voltage that induces the charge $Q_0 \simeq e$ on the junction capacitance. The two states $n = 0, 1$ are nearly-degenerate and the junction behaves effectively as a macroscopic two-level system.

Thus, the appropriately biased small Josephson tunnel junction is a macroscopic two-level system, with the two states represented by the position of a single Cooper pair on the left or right electrode of the junction. In principle, this system can be used as a q-bit of the quantum logic gates. However, if q-bits are represented with single junctions, neither the tunneling amplitude $E_J/2$ nor the coupling strength of the two q-bits which is determined by the coupling capacitance between the junction electrodes can be modulated in time as required by the design of the adiabatic CN gate. In particular, to realize adiabatic dynamics it should be possible to suppress both the tunneling amplitude and interaction strength to zero between the active cycles of the gate operation. This problem can be circumvented if q-bits are represented not with individual junctions but with the one-dimensional arrays of junctions. In an array, the tunneling amplitude Ω between the two islands of the array can be effectively modulated by the gate voltages applied to the islands of the array, and the interaction energy η of charges in the array decreases exponentially with the distance between them.

To make a q-bit out of a uniform array, all islands should have individual gate electrodes supplying the gate voltages V_j (Fig. 3a,b), and two internal islands of the array should be biased with the voltages $\pm e/C_t$, where $C_t = (C_0^2 + 4CC_0)^{1/2}$ is the total capacitance of an internal island in the array – see, e.g., [14], and C , C_0 are, respectively, the junction capacitance, and the capacitance between each island and its gate electrode (Fig. 3b). These voltages induce the charges e and $-e$ on the two islands, so that the two charge configurations of the array: one with no Cooper pair transferred across any junction and another one with a Cooper pair transferred between the two biased islands, from e to $-e$, have the same energy. This means that if the bias conditions do not deviate strongly from these conditions, all other charge configurations of the array have much larger energies and the array dynamics is equivalent to the two-state dynamics that can be described in terms of the tunneling of a single Cooper pair between the two islands. In this regime the array can be viewed as a q-bit with the two positions of the Cooper pair on one or another island representing the two states of the computational basis of this q-bit.

If the two islands containing the q-bit states are separated by m junctions, the amplitude of tunneling Ω between them depends exponentially on the separation m . The dominant contribution to Ω comes from the process in which the Cooper pair is transferred sequentially through the junctions separating the islands, and can be written as:

$$\Omega = \frac{E_J}{2} \prod_{k=1}^{m-1} \frac{E_J}{2E_k}, \quad (5)$$

where E_k are the energies of the intermediate charge configurations resulting from the Cooper pair transfer through the first k junctions. These energies are controlled by the gate voltages applied to the intermediate islands.

The most important feature of the Cooper pair states forming q-bit basis is that they can be moved along the array by the adiabatic level-crossing transitions similar to those discussed above. A Cooper pair is transferred between the two adjacent islands when a gate voltage of the initially occupied island

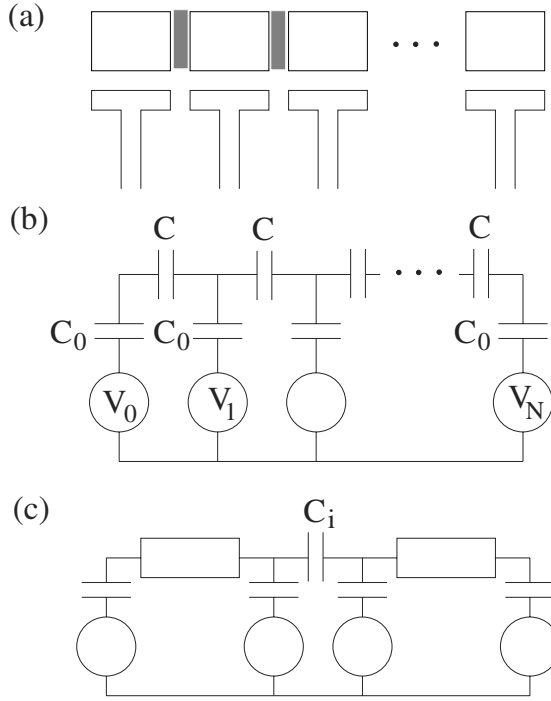


Fig. 3. Schematic layout (a) and equivalent electrostatic circuit (b) of an array of small Josephson junctions representing one q-bit. (c) The controlled-NOT gate obtained by coupling of the two arrays.

is increased/decreased while the gate voltage of the neighboring island is decreased/increased adiabatically past e/C_t . The adjacent islands are coupled by the tunneling amplitude $E_J/2$, and the Cooper pair is transferred with the probability exponentially approaching one if the rate of change of the gate voltages is small on the scale of this amplitude. Similar manipulation of the gate voltages also shifts the empty state of the q-bit by one island. In this way it is possible to move the q-bit states around, either shifting both states along the array, or changing the separation m between the two states.

This dynamics is analogous to the one used in the so-called single-electron [16] and single Cooper pair [17] pump, or single-electron parametron [18]. It allows to implement the general scheme of the adiabatic CN gate with the two coupled arrays representing the two q-bits of the gate (Fig. 3c). As a first step of the gate operation, the q-bit states in both arrays are moved towards the ends of the arrays where they can interact via the coupling capacitance C_i . The states of the controlled q-bit in the first array have sufficiently large separation m so that their tunnel coupling Ω_1 is negligible. By contrast, the states of the target q-bit in the second array are put on the adjacent islands in order to maximize their tunnel coupling, $\Omega_2 = E_J/2$. Then a pulse of the bias voltage is applied to

the first junction of the target q-bit array. If the control q-bit is in the “1” state, a Cooper pair occupies the island of the first array closest to the second array and creates additional potential drop δV across the junction of the target q-bit:

$$\delta V = \frac{8eC_i}{(C_0 + C_t + 2C)(C_0 + C_t + 4C_i)}. \quad (6)$$

In this case the bias pulse drives the target q-bit through the level-crossing point so that the occupation probabilities of its states are interchanged. When the control q-bit is in the “0” state, the Cooper pair of this q-bit is inside the array and does not produce extra voltage across the target q-bit junction, which then does not reach the level-crossing point, and the occupation probabilities of its states remain the same. During the last step of the gate operation it is returned to its initial configuration, i.e., the separation of the states of the target q-bit is increased to suppress the tunnel amplitude Ω_2 to zero, and the states of the both q-bits are shifted inside the arrays. Then the interaction of the q-bit states becomes negligible due to screening by the gate electrodes, which is known to lead to the exponential suppression of the interaction energy η between two Cooper pairs separated by m junctions of one array [14]:

$$\eta = \frac{(2e)^2}{C_t} \lambda^m, \quad \lambda = \frac{2C}{2C + C_0 + C_t}. \quad (7)$$

This implementation of the CN quantum gate can only be practical if it is stable against deviations of the real gate structure from the idealized model used above. Such deviations are fundamentally unavoidable in all macroscopic realizations of quantum gates. For instance, the real electrostatics of the Josephson junction gate is much more complicated than the model characterized by the two nearest-neighbor capacitances C and C_0 . It involves full capacitance matrix C_{ij} in which even remote islands interact with each other, and should also describe small fluctuations of the nearest-neighbor capacitances around their average values. An important advantage of the adiabatic approach is that these complications can be compensated for by the adjustment of the bias voltages and do not change qualitatively the gate dynamics. Indeed, the adiabatic transfer of a Cooper pair depends only on the resonance condition that the energies of all Cooper pair states along the array are the same, which ensures correct transfer of the occupation probabilities of the gate states. The bias voltages can always be tuned to satisfy the resonance condition regardless of the form of the capacitance matrix. A practical proof of this statement is provided by the experimentally demonstrated operation of a similar system, single-electron pump, with accuracy better than 10^{-6} [19].

The only instance when the gate dynamic relies heavily on the simplified model of the array electrostatic is in the assumption of the exponential screening of the electrostatic interaction inside the array. In the realistic model of electrostatics, interaction at large distances depends on the external environment of the array. The exponential screening of the interaction can still be obtained even in this case, but requires that the array is placed between the two conducting ground planes.

These considerations show that dynamics of the occupation probabilities of the gate states is indeed insensitive to the weak disorder in the gate parameters. However, the proper dynamics of the system as quantum logic gate requires also that the phases of the occupation amplitudes accumulated during the gate operation are all equal modulo 2π . In this respect, fluctuations of the junction parameters do present a problem since they make the dynamic phases of the gate states unpredictable. This problem can be resolved if the disorder in the parameters is static on a sufficiently long time scale. In this case, the phases can be measured and compensated for by the fine-tuning of the gate voltages.

In order to measure the phases, we need to transform them into the occupation probabilities of the gate states which in their turn can be measured with a single-electron electrometer (see, e.g., [20], Chapter 9). An electrometer measures an average charge of the island and therefore gives information about the occupation probabilities of the gate state, but is insensitive to the phase of the occupation amplitudes. Suppose that as a result of a prior measurements, we know that the occupation probabilities of the two q-bit states are p_1 and p_2 . The two states are decoupled (the corresponding tunneling amplitude Ω is zero) and their energies are equal, so that there is some stationary phase difference φ between their occupation amplitudes. The phase φ can be transformed into the occupation probability by rotation \hat{U} of the q-bit states in the Hilbert state, $\hat{U} = \exp\{i\pi\sigma_x/4\}$. This rotation is achieved if the barrier between the states is reduced temporarily in such a way that

$$\int dt \Omega(t) = \pi\hbar/4.$$

The resulting occupation probabilities

$$q_{1,2} = 1/2 \mp (p_1 p_2)^{1/2} \sin \varphi,$$

depend on the phase φ , and by measuring them we can measure φ . After the phase is known it can be compensated for by adding an extra voltage pulses at the end of the gate operation. With this fine-tuning, the gate dynamics becomes effectively independent of the static disorder in the gate parameters.

4 Decoherence Rate and Parameter Estimates

The above discussion assumes that the energy relaxation and associated with it time-dependent fluctuations of the phase are negligible. There are several dissipation and dephasing mechanisms in the Josephson tunnel junction systems. Some of them are well understood and can be controlled within certain limits. One of these mechanisms is the quasiparticle tunneling. In general, it coexists with the Cooper pair tunneling and makes junction dynamics irreversible. However, if both the temperature T and charging energy E_C of the junctions are much smaller than the superconducting energy gap Δ , the quasiparticle tunneling is suppressed by the parity effects [21,22,23] to a level where it can be negligible on

the macroscopic time scales [24,25]. Another dissipation mechanism is coupling to the electromagnetic excitations supported by the system of superconducting electrodes. A Cooper pair oscillating between the two islands creates oscillating currents in the islands and electric fields around the islands which couple to these modes. The power P lost to electromagnetic modes depends on the specific geometry of the islands and connecting them tunnel junctions. Part of the losses comes from the direct dipole radiation from the junctions and can be estimated as radiation of a dipole of length equal to the length d of the junction electrodes:

$$P_d \simeq \frac{e^2 \omega^4 d^2}{4\pi \epsilon_0 c^3}. \quad (8)$$

The radiated power is not exponentially small, nevertheless it decreases sufficiently rapidly with decreasing ratio of the island size to the radiation wavelength $\lambda \simeq c/\omega$ at frequency $\omega \simeq E_J/\hbar$. Therefore, to keep this type of radiation losses small the islands of the junction arrays should be much smaller than the wavelength at frequency E_J/\hbar , the condition that is always satisfied in small junctions.

The crucial contribution to radiation losses comes from the coupling to electromagnetic modes supported by essentially “infinite” external gate electrodes supplying bias voltages to the islands. In the relevant regime with $C_0 \ll C$, the power dissipated into these modes can be estimated in terms of the wave impedance ρ of the gate electrodes as

$$P_l \simeq \left(\frac{eC_0}{C}\right)^2 \omega^2 \rho. \quad (9)$$

We see that this dissipation mechanism limits the magnitude of the island capacitance to the gate electrodes C_0 . In the simple model of the gate electrostatics, C_0 determines also the number of islands of the junction array that are polarized by a single Cooper pair, and restriction on C_0 translates into the limitation on how small the number of junctions in the arrays can be. If however, one introduces ground planes which give rise to extra stray capacitances of the array islands, these two limitations becomes uncoupled. In any case, for realistic values of the parameters (see the estimates below) the losses (9) in the external electrodes should give the dominant contribution to decoherence for the Cooper pair tunneling.

Besides these “controllable” mechanisms of dissipation that depend on the gate geometry, the Cooper pair tunneling in the junction arrays is affected also by the “internal” dissipation in all elements of the arrays. The most important source of noise and dissipation of this kind is the $1/f$ charge noise in the insulators surrounding the junctions: substrate and tunnel barriers. The strength of the noise is material dependent and can not be estimated from first principles. Experimentally, characteristic time scale of the charge noise varies from millisecond range [26] to seconds and hours [27], and is much longer than characteristic time of the Cooper pair tunneling \hbar/E_J which determines the rate of the gate operation. Therefore, the gate can go through the large number of cycles

of operation before the decoherence due to the charge noise starts to affect its dynamics.

In conclusion, we summarize the conditions that should be satisfied by junction arrays in order to operate as quantum logic gates. The first set of conditions is given by the following string of inequalities:

$$T \ll E_J \ll E_C \ll \Delta. \quad (10)$$

The two limiting energy scales in this relations, temperature T and energy gap Δ , are practically constrained by the available refrigeration technology and superconducting materials. The lower limit is set by the typical electron temperature attainable in experiments with the dilution refrigerator and is on the order of 30 mK. The upper limit can not be much larger than the energy gap of niobium, or its compounds, i.e., about 20 K. The ratio of the Josephson coupling energy E_J to the charging energy E_C can not be varied arbitrarily because of the technological limitations on the critical current density that can be obtained while preserving the quality of the tunnel junction. Conditions (10) are satisfied if we take, for instance, $E_J \simeq 1$ K, and $E_C \simeq 3$ K. This value of E_C corresponds to the junction capacitance $C \simeq 0.5$ fF, which for a typical specific capacitance of a tunnel junction, 0.1 pF/ μm^2 , requires the junction area of about 70×70 nm². With this area, the cited E_J value corresponds to the critical current density j_c about 10 $\mu\text{A}/\mu\text{m}^2$, and the total critical current $I_c = 2eE_J/\hbar \simeq 50$ nA. Experimentally, this value of j_c is within the range of current densities that can be achieved without the degradation of the tunnel junction quality [28].

Another condition on the junction array as a CN gate is that the number N of junctions in it is much larger than its screening length:

$$N \gg (C/C^*)^{1/2}.$$

Here C^* is the total stray capacitance of the array islands which include capacitance C_0 to the gate electrodes and capacitance to the ground planes. This condition does not represent a serious obstacle to realization of a CN gate. Specific values of N and C^* are dictated by the convenience of fabrication of either longer arrays or larger capacitances to the ground.

The most difficult is the condition that the probability α of the decoherence-induced error during one cycle of the gate operation is small. Estimating the period of this cycle roughly as \hbar/E_J we obtain from eq. (9) that the lower bound on α is:

$$\alpha \simeq \left(\frac{C_0}{C}\right)^2 \frac{e^2 \rho}{\hbar}. \quad (11)$$

The values of parameters that are typical for existing experiments (in which no effort was made to decrease α) are $C_0/C \simeq 0.1$, and $\rho \simeq 300$ Ohm [29]. (The latter value corresponds to a narrow, about $1\mu\text{m}$, electrode.) In this case $\alpha \simeq 10^{-3}$. The error probability can be substantially reduced by making coupling capacitance C_0 smaller, and gate electrodes wider thus decreasing ρ . Although only experiments can tell what is the limit to decrease in decoherence rate, it is

reasonable to expect that α can be reduced further by a few orders of magnitude to a value about 10^{-6} .

The author thanks J.P. Pekola and K.-A. Suominen for critical reading of the manuscript. This work was supported by ONR.

References

1. A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996); C.H. Bennett, Physics Today, October 1995, p. 24; D. DiVincenzo, Science **270**, 255 (1995).
2. P.W. Shor, in: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, (IEEE Computer Society, Los Alamitos, CA, 1994), p. 124.
3. L.K. Grover, *A fast quantum mechanical algorithm for database search*, quant-ph/9605043.
4. S. Haroche and J.-M. Raimond, Physics Today, August 1996, p. 51.
5. J.I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).
6. C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano, and D.J. Wineland, Phys. Rev. Lett. **75**, 4714 (1995).
7. Q.A. Turchette, C.J. Hood, W. Lange, H. Mabuchi, and H.J. Kimble, Phys. Rev. Lett. **75**, 4710 (1995).
8. A. Barenko, D. Deutch, A. Ekert, and R. Jozsa, Phys. Rev. Lett. **74**, 4083 (1995).
9. D. Loss and D. DiVincenzo, cond-mat/9701055.
10. N.A. Gershenfeld and I.L. Chuang, Science **275**, 350 (1997).
11. M.F. Bosco, A.M. Herr, and M.J. Feldman, IEEE Trans. Appl. Supercond. **7**, 3638 (1997).
12. C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W.K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
13. D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).
14. D.V. Averin and K.K. Likharev, in: *Mesoscopic Phenomena in Solids*, ed. by B.L. Altshuler et al. (Elsevier, Amsterdam, 1991), p. 173.
15. M. Tinkham, *Introduction to Superconductivity*, (McGraw-Hill, New York, 1996), Chapters 6,7.
16. H. Pothier, P. Lafarge, P.F. Orfila, C. Urbina, D. Esteve, and M.H. Devoret, Physica B **169**, 573 (1991).
17. L.G. Geerligs, S.M. Verbrugh, P. Hadley, J.E. Mooij, H. Pothier, P. Lafarge, C. Urbina, D. Esteve, and M.H. Devoret, Zs. Phys. B **85**, 349 (1991).
18. K.K. Likharev and A.N. Korotkov, Science **273**, 763 (1996).
19. J.M. Martinis, M. Nahum, and H.D. Jensen, Phys. Rev. Lett. **72**, 94 (1994).
20. *“Single Charge Tunneling”*, ed. by H. Grabert and M. Devoret (Plenum, NY, 1992).
21. D.V. Averin and Yu.V. Nazarov, Physica B **203**, 310 (1994).
22. J.G. Lu, J.M. Hergenrother, and M. Tinkham, Phys. Rev. B **53**, 3543 (1996).
23. Y. Nakamura, C.D. Chen, and J.S. Tsai, Czechoslovak Journal of Physics **46**, Suppl. 6, 3339 (1996).
24. T.M. Eiles, J.M. Martinis, and M.H. Devoret, Phys. Rev. Lett. **70**, 1862 (1993).
25. J.E. Lukens, P.D. Dresselhaus, S. Han, L. Ji, K.K. Likharev, and W. Zheng, Physica B **203**, 354 (1994).
26. J.M. Hergenrother, M.T. Tuominen, T.S. Tighe, and M. Tinkham, IEEE Trans. Appl. Supercond. **3**, 1980 (1993).

27. P.D. Dresselhaus, L. Ji, S. Han, J.E. Lukens, and K.K. Likharev, Phys. Rev. Lett. **72**, 3226 (1994).
28. A.W. Kleinsasser, R.E. Miller, W.H. Mallison, and G.B. Arnold, Phys. Rev. Lett. **72**, 1738 (1994).
29. J.P. Kauppinen and J.P. Pekola, Phys. Rev. Lett. **77**, 3889 (1996).

Trapped Ion Quantum Computer Research at Los Alamos

D.F.V. James, M.S. Gulley, M.H. Holzscheiter, R.J. Hughes, P.G. Kwiat,
S.K. Lamoreaux, C.G. Peterson, V.D. Sandberg, M.M. Schauer,
C.M. Simmons, D. Tupa, P.Z. Wang, and A.G. White

Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Abstract. We briefly review the development and theory of an experiment to investigate quantum computation with trapped calcium ions. The ion trap, laser and ion requirements are determined, and the parameters required for simple quantum logic operations are described.
(LAUR 98-314)

1 Introduction

In the last 15 years various authors have considered the generalization of information theory concepts to allow the representation of information by quantum systems. The introduction into computation of *quantum mechanical* concepts, in particular the superposition principle, opened up the possibility of new capabilities, such as quantum cryptography [1], that have no classical counterparts. One of the most interesting of these new ideas is quantum computation, first proposed by Benioff [2]. Feynman [3] suggested that quantum computation might be more powerful than classical computation, a notion which gained further credence through the work of Deutsch [4]. However, until quite recently quantum computation was an essentially academic endeavor because there were no quantum algorithms that exploited this power to solve useful computational problems, and because no realistic technology capable of performing quantum computations had been envisioned. This changed in 1994 when Shor discovered quantum algorithms for efficient solution of integer factorization and the discrete logarithm problem [5,6], two problems that are at the heart of the security of much of modern public key cryptography [7]. Later that same year Cirac and Zoller proposed that quantum computational hardware could be realized using known techniques in the laser manipulation of trapped ions [8]. Since then interest in quantum computation has grown dramatically, and remarkable progress has been made: a single quantum logic gate has been demonstrated with trapped ions [9]; quantum error correction schemes have been invented [10,11]; several alternative technological proposals have been made [21,22,23,24,25,26] and quantum algorithms for solving new problems have been discovered [16,17,18,19]. In this paper we will review our development of an experiment to investigate the potential of quantum computation using trapped calcium ions [15].

The three essential requirements for quantum computational hardware are: (1) the ability to isolate a set of two-level quantum systems from the environment for long enough to maintain coherence throughout the computation, while at the same time being able to interact with the systems strongly enough to manipulate them into an arbitrary quantum state; (2) a mechanism for performing quantum logic operations: in other words a “quantum bus channel” connecting the various two-level systems in a quantum mechanical manner; and (3) a method for reading out the quantum state of the system at the end of the calculation.

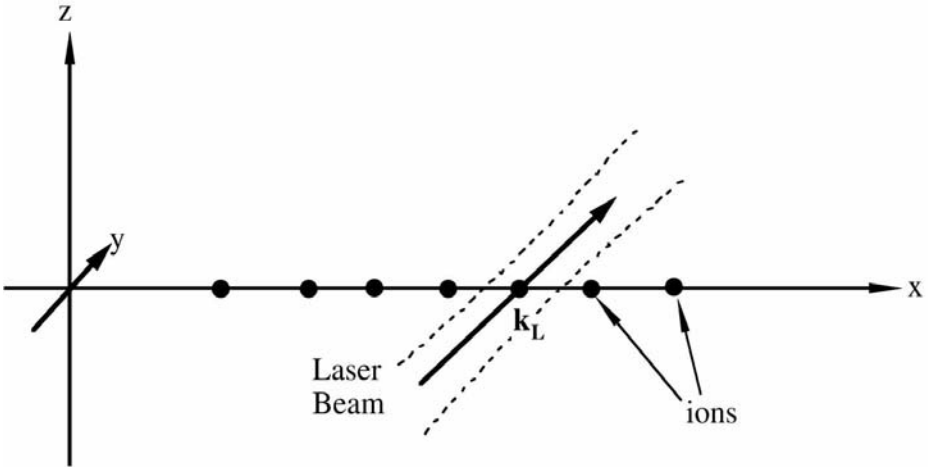


Fig. 1. A schematic illustration of an idealized laser-ion interaction system; \mathbf{k}_L is the wavevector of the single addressing laser.

All three of these requirements are in principle met by the cold trapped ion quantum computer. In this scheme each qubit consists of two internal levels of an ion trapped in a linear configuration. In order to perform the required logic gates, a third atomic state known as the auxiliary level is required. The quantum bus channel is realized using the phonon modes of the ions' collective oscillations. These quantum systems may be manipulated using precisely controlled laser pulses. Two distinct types of laser pulse are required: “V” type pulses, which only interact with the internal states of individual ions, and “U” type pulses which interact with both the internal states and the external vibrational degrees of freedom of the ions. These interactions can be realized using Rabi flipping induced by either a single laser or Raman (two laser) scheme (Fig.2). Readout is performed by using quantum jumps. This scheme was originally proposed by Cirac and Zoller in 1994 [8], and was used to demonstrate a CNOT gate shortly afterwards [9].

As we can only give the briefest of description of the principles of quantum computation using cold trapped ions, the reader is recommended to peruse

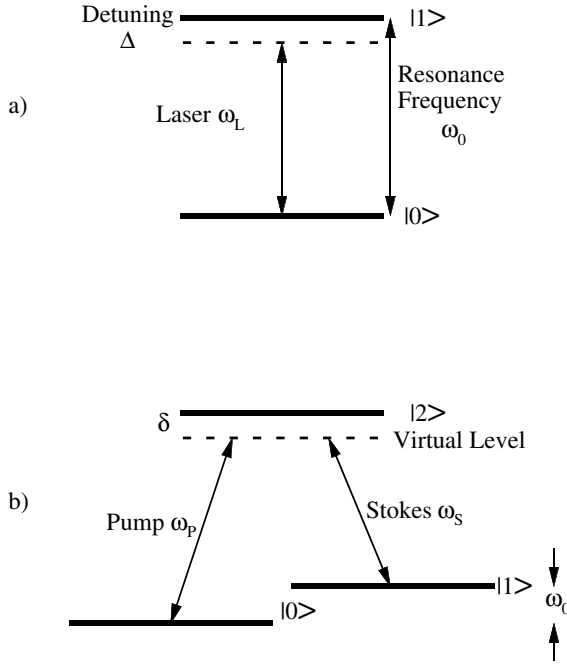


Fig. 2. A schematic illustration of (a) single laser and (b) Raman qubit addressing and control techniques.

the more detailed descriptions which can be found elsewhere [12,13,14,15]. In this paper we intend to focus on the experimental issues involved in building a trapped ion quantum computer.

2 Choice of Ion

There are three requirements which the species of ion chosen for the qubits of an ion trap quantum computer must satisfy:

1. If we use the single laser scheme, the ions must have a level that is sufficiently long-lived to allow some computation to take place; this level can also be used for sideband cooling.
2. the ions must have a suitable dipole-allowed transition for Doppler cooling, quantum jump readout and for Raman transitions (if we chose to use two sub-levels of the ground state to form the qubit);
3. These transitions must be at wavelengths compatible with current laser technology.

Various ions used in atomic frequency standards work satisfy the first requirement. Of these ions, Ca^+ offers the advantages of transitions that can be accessed with titanium-sapphire or diode lasers and a reasonably long-lived metastable state. The relevant energy levels of the $A = 40$ isotope are shown in fig.3.

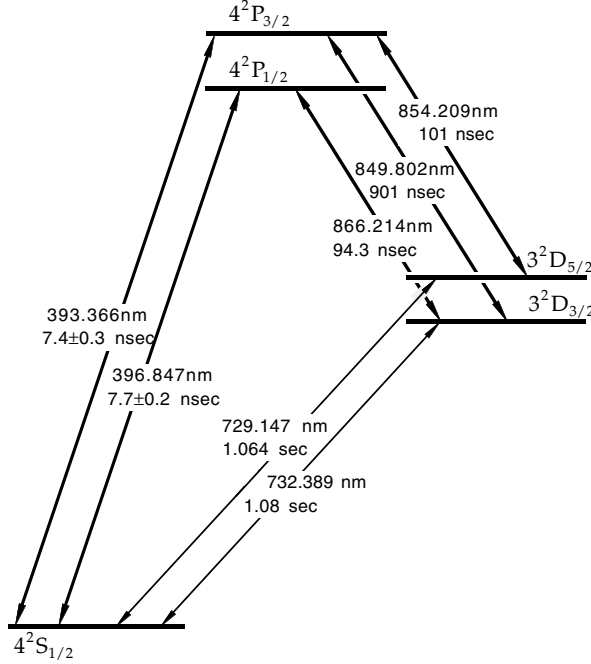


Fig. 3. The lowest energy levels of $^{40}\text{Ca}^+$ ions, with transition wavelengths and lifetimes listed.

The dipole-allowed transition from the $4^2S_{1/2}$ ground state to the $4^2P_{1/2}$ level with a wavelength of 397 nm can be used for Doppler cooling and quantum jump readout; The 732 nm electric quadrupole transition from the $4^2S_{1/2}$ ground state to the $3^2D_{3/2}$ metastable level (lifetime $\approx 1.08\text{sec.}$) is suitable for sideband cooling. In the single laser computation scheme, the qubits and auxiliary level can be chosen as the electronic states

$$\begin{aligned} |0\rangle &= |4^2S_{1/2}, M_j = 1/2\rangle, \\ |1\rangle &= |3^2D_{5/2}, M_j = 3/2\rangle, \\ |aux\rangle &= |3^2D_{5/2}, M_j = -1/2\rangle. \end{aligned}$$

This ion can also be used for Raman type qubits, with the two Zeeman sublevels of the $4^2S_{1/2}$ ground state forming the two qubit states $|0\rangle$ and $|1\rangle$, with one of the sublevels of the $4^2P_{1/2}$ level being the upper level $|2\rangle$. A magnetic field of 200 Gauss should be sufficient to split these two levels so that they can be resolved by the lasers. The pump and Stokes beams would be formed by splitting a 397nm laser into two, and shifting the frequency of one with respect to the other by means of an acousto-optic or electro-optic modulator. This arrangement has a great advantage in that any fluctuations in the phase of the original 397nm

laser will be passed on to both the pump and Stokes beams, and will therefore be canceled out, because the dynamics is only sensitive to the difference between the pump and Stokes phases. One problem in realizing the Raman scheme in Ca^+ is the absence of a third level in the ground state that can act as the auxiliary state $|aux\rangle$ required for execution of quantum gates. This difficulty could be removed by using the alternative scheme for quantum logic recently proposed by Monroe *et al.* [27]; alternatively, one could use an isotope of Ca^+ which has non-zero nuclear spin, thereby giving several more sublevels in the ground state due to the hyperfine interaction; other possibilities that have been suggested for an auxiliary state with $^{40}Ca^+$ in the Raman scheme are to use a state of a phonon mode other than the CM mode [28] or one of the sublevels of the 3^2D doublet [29].

3 The Radio Frequency Ion Trap

Radio-frequency (RF) quadrupole traps, also named “Paul traps” after their inventor, have been used for many years to confine electrically charged particles [30] (for an introduction to the theory of ion traps, see refs. [31,32]). The classic design of such a Paul trap has a ring electrode with endcap electrodes above and below, with the ions confined to the enclosed volume. A single ion can be located precisely at the center of the trap where the amplitude of the RF field is zero. But when several ions are placed into this trapping field, their Coulomb repulsion forces them apart and into regions where they are subjected to heating by the RF field. For this reason in our experiment ions are confined in a linear RF quadrupole trap [15]. Radial confinement is achieved by a quadrupole RF field provided by four 1 mm diameter rods in a rectangular arrangement. Axial confinement is provided by DC voltages applied to conical endcaps at either end of the RF structure; the endcap separation is 10 mm. The design of the trap used in these experiments is shown in diagrammatically in Fig.4.

The main concerns for the design are to provide sufficient radial confinement to assure that the ions form a string on the trap axis after Doppler cooling; to minimize the coupling between the radial and axial degrees of freedom by producing radial oscillation frequencies significantly greater than the axial oscillation frequencies; to produce high enough axial frequencies to allow the use of sideband cooling [33]; and to provide sufficient spatial separation to allow individual ions to be addressed with laser beams.

4 Laser Systems

The relevant optical transitions for Ca^+ ions are shown in Fig.5. There are four different optical processes employed in the quantum computer; each places specific demands on the laser system.

The first stage is to cool a small number of ions to their Doppler limit in the ion trap, as shown in Fig.5a. This requires a beam at 397 nm, the $4^2S_{1/2} - 4^2P_{1/2}$

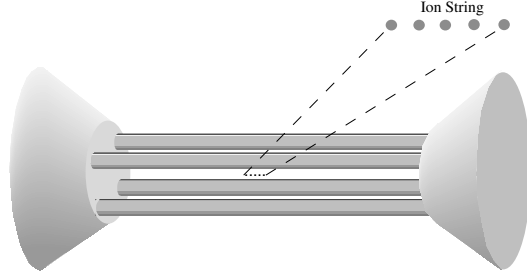


Fig. 4. Side view diagram of the linear RF trap used to confine Ca^+ ions in these experiments. The endcap separation is 10 mm and the gap between the RF rods is 1.7 mm.

resonant transition. Tuning the laser to the red of the transition causes the ions to be slowed by the optical molasses technique [34]. In this procedure, a laser beam with a frequency slightly less than that of the resonant transition of an ion is used to reduce its kinetic energy. Owing to the Doppler shift of the photon frequency, ions preferentially absorb photons that oppose their motion, whereas they re-emit photons in all directions, resulting in a net reduction in momentum along the direction of the laser beam. Having carefully selected the trap parameters, many cycles of absorption and re-emission will bring the system to the Lamb-Dicke regime, leaving the ions in a string-of-pearls geometry. We have recently found ion crystals of up to five Ca^+ ions.

In order to Doppler cool the ions, the demands on the power and linewidth of the 397 nm laser are modest. The saturation intensity of Ca^+ ions is $\sim 10 \text{ mW/cm}^2$, and the laser linewidth must be less than $\sim 10 \text{ MHz}$. An opto-galvanic signal obtained with a hollow cathode lamp suffices to set the frequency. We use a Titanium:Sapphire (Ti:Sapphire) laser (Coherent CR 899-21) with an internal frequency doubling crystal to produce the 397 nm light.

During the Doppler cooling, the ions may decay from the $4^2P_{1/2}$ state to the $3^2D_{3/2}$ state, whose lifetime is $\sim 1 \text{ sec}$. To empty this metastable state, we use a second Ti:Sapphire laser at 866 nm.

Once the string of ions is Doppler cooled to the Lamb-Dicke regime, the second stage of optical cooling, sideband cooling, will be used to reduce the collective motion of the string of ions to its lowest vibrational level [35], illustrated in Fig.5b. In this regime, a narrow optical transition, such as the 732 nm $4^2S_{1/2} - 3^2D_{3/2}$ dipole forbidden transition, develops sidebands above and below the central frequency by the vibrational frequencies of the ions. The sidebands closest to the unperturbed frequency correspond to the CM vibrational motion. If ω_0 is the optical transition frequency and ω_x the frequency of the CM vibrational motion, the phonon number is increased by one, unchanged, or decreased by one if an ion absorbs a photon of frequency $\omega_0 + \omega_x$, ω_0 or $\omega_0 - \omega_x$,

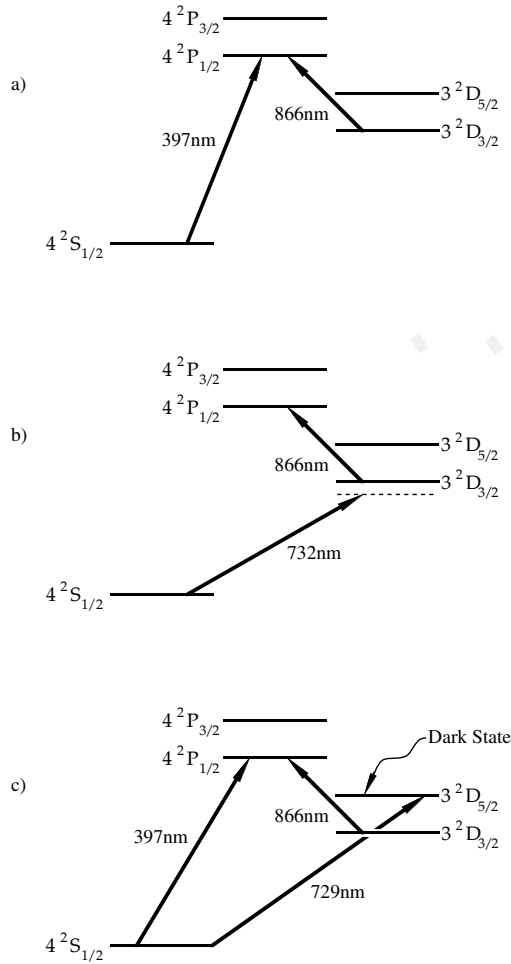


Fig. 5. Different transitions between the levels of Ca^+ ions required for (a) Doppler cooling, (b) Resolved sideband cooling and (c) quantum logic operations and readout. The single laser addressing technique has been assumed.

respectively. Thus, sideband cooling is accomplished by optically cooling the string of ions with a laser tuned to $\omega_0 - \omega_x$.

The need to resolve the sidebands of the transition implies a much more stringent requirement for the laser linewidth; it must be well below the CM mode vibrational frequency of $\sim (2\pi) \times 200$ kHz. The laser power must also be greater in order to pump the forbidden transition. We plan to use a Ti:Sapphire laser locked to a reference cavity to meet the required linewidth and power. At first glance it would seem that, with a metastable level with a lifetime of 1s, no more than 1 phonon per second could be removed from a trapped ion. A second laser

at 866 nm is used to couple the $4^2P_{1/2}$ state to the $3^2D_{3/2}$ state to reduce the effective lifetime of the D state and allow faster cooling times. The transitions required for realization of quantum logic gates and for readout, discussed in detail in sections 5.2 and 5.3, are shown in Fig.5c. These can be performed with the same lasers used in the Doppler and sideband cooling procedures.

There are two other considerations concerning the laser systems for quantum computation which should be mentioned. To reduce the total complexity of the completed system, we are developing diode lasers and a frequency doubling cavity to handle the Doppler cooling and quantum jump read out. Also complex quantum computations would require that the laser on the $4^2S_{1/2} - 3^2P_{5/2}$ computation transition have a coherence time as long as the computation time. This may necessitate using qubits bridged by Raman transitions as discussed above, which eliminates the errors caused by the phase drift of the laser.

5 Qubit Addressing Optics

In order for the Ca^+ ion qubits to be useful for actual calculations, it will be necessary to address the ions in a very controlled fashion. Our optical system for qubit addressing is shown schematically in fig. 6.

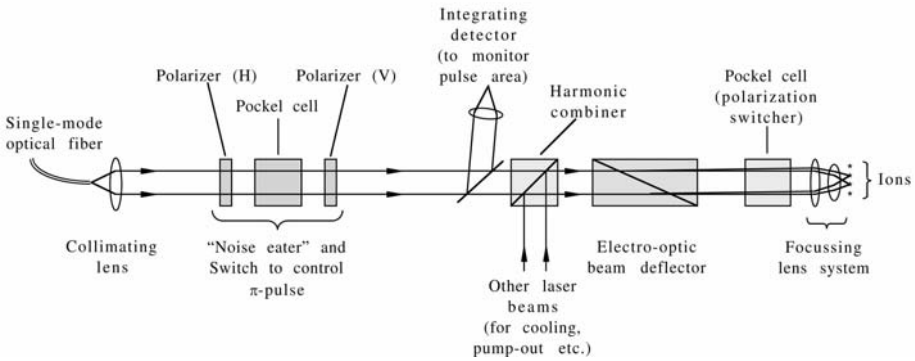


Fig. 6. Illustration of the laser beam control optics system.

There are two aspects to be considered in the design of such a system: the precise interactions with a single ion; and an arrangement for switching between different ions in the string. In addition to the obvious constraints on laser frequency and polarization, the primary consideration for making exact π - or 2π -pulses is control of the area (over time) of the driving light field pulse. The first step toward this is to stabilize the intensity of the laser, as can be done to better than 0.1%, using a standard “noise-eater”. Such a device typically consists of an electro-optical polarization rotator located between two polarizers; the output of a fast detector monitoring part of the transmitted beam is used in a feedback

circuit to adjust the degree of polarization rotation, and thus the intensity of the transmitted light. Switching the light beam on and off can be performed with a similar (or even the same) device. Because such switches can possess rise/fall times on the scale of nanoseconds, it should be possible to readily control the area under the pulse to within $\sim 0.1\%$, simply by accurately determining the width of the pulse. A more elaborate scheme would involve an integrating detector, which would monitor the actual integrated energy under the pulse, shutting the pulse off when the desired value is obtained.

Once the controls for addressing a single ion are decided, the means for switching between ions must be considered. Any system for achieving this must be fast, reproducible, display very precise aiming and low “crosstalk” (i.e. overlap of the focal spot onto more than one ion), and be as simple as possible. In particular, it is desirable to be able to switch between different ions in the string in a time short compared to the time required to complete a given π -pulse on one ion. This tends to discount any sort of mechanical scanning system. Acousto-optic deflectors, which are often used for similar purposes, may be made fast enough, but introduce unwanted frequency shifts on the deviated beams. As a tentative solution, we propose to use an electro-optic beam deflector, basically a prism whose index of refraction, and consequently whose deflection angle, is varied slightly by applying a high voltage across the material; typical switching times for these devices is 10 nanoseconds, adequate for our purposes. One such device produces a maximum deflection of ± 9 mrad, for a ± 3000 V input. The associated maximum number of resolvable spots (using the Rayleigh criterion) is of order 100, implying that ~ 20 ions could be comfortably resolved with negligible crosstalk.

After the inter-ion spacing has been determined, i.e., by the trap frequencies, the crosstalk specification determines the maximum spot size of the addressing beam. For example, for an ion spacing of $20\text{ }\mu\text{m}$, any spot size (defined here as the $1/e^2$ diameter) less than $21.6\text{ }\mu\text{m}$ will yield a crosstalk of less than 0.1% , assuming a purely Gaussian intensity distribution (a good approximation if the light is delivered from a single-mode optical fiber, or through an appropriate spatial filter). In practice, scattering and other experimental realities will increase this size, so that it is prudent to aim for a somewhat smaller spot size, e.g. $10\text{ }\mu\text{m}$. One consideration when such small spot sizes are required is the effect of lens aberrations, especially since the spot must remain small regardless of which ion it is deflected on. Employing standard ray-trace methods, we have found that the blurring effects of aberrations can be reduced if a doublet/meniscus lens combination is used (assuming an input beam size of 3mm , and an effective focal length of 30mm). A further complication is that, in order to add or remove phonons from the system, the addressing beams must have a component along the longitudinal axis of the trap. The addressing optics must accommodate a tilted line of focus, otherwise the intensity at each ion would be markedly different, and the crosstalk for the outermost ions would become unacceptable. According to ray-trace calculations, adding a simple wedge (of $\sim 2^\circ$) solves the problem and this has been confirmed by measurements using a mock system.

Depending on the exact level scheme being considered, it may be necessary to vary the polarization of the light. Because the electro-optic deflector requires a specific linear polarization, any polarization-control elements should be placed after the deflector. The final result is a highly directional, tightly-focused beam with controllable polarization and intensity.

6 Imaging System

In order to determine the ions' locations and to readout the result of the quantum computations, an imaging system is required. Our current imaging system consists of two lenses, one of which is mounted inside the vacuum chamber, and a video camera coupled to a dual-stage micro-channel plate (MCP) image intensifier. The first lens with focal length 15 mm collects the light emitted from the central trap region with a solid angle of approximately 0.25 sr. The image is relayed through a 110mm/f2 commercial camera lens to the front plate of the MCP. This set-up produces a magnification of 7.5 at the input of the MCP. The input of the 110 mm lens is fitted with a 400nm narrow band filter to reduce background from the IR laser and from light emanating from the hot calcium oven and the electron gun filament.

The dual plate intensifier is operated at maximum gain for the highest possible sensitivity. This allows us to read out the camera at normal video rate of 30 frames s^{-1} into a data acquisition computer. Averaging and integrating of the signal over a given time period can then be undertaken by software. We find this arrangement extremely useful in enabling us to observe changes of the cloud size or the intensity of the fluorescence with changes of external parameters like trapping potential, laser frequency, laser amplitude, etc. in real time.

The spatial resolution of the system is limited by the active diameter of individual channels of the MCP of approximately 12 μm . Since the gain is run at its maximum value cross talk between adjacent channels in the transition between the first and second stage is to be expected. This results in the requirement that two incoming photons can only be resolved when they are separated at the input of the MCP by at least two channels, i.e. by 36 μm in our case. With the magnification of the optical system of 7.5 this translates into a minimum separation of two ions to be resolved of 5 μm , which is well below the separation of ions in the axial well of about 25 μm expected in our experiment.

7 Summary

It is our contention that currently the ion trap proposal for realizing a practical quantum computer offers the best chance of long term success. This in no way is intended to trivialize research into the other proposals: in any of these schemes technological advances may at some stage lead to a breakthrough. In particular, Nuclear Magnetic Resonance does seem to be a relatively straightforward way in which to achieve systems containing a few qubits. However, of the technologies which have so far been used to demonstrate experimental logic gates, ion traps

seem to offer the least number of technological problems for scaling up to 10's or even 100's of qubits.

In this paper we have described in some detail the experiment we are currently developing to investigate the feasibility of cold trapped ion quantum computation. We should emphasize that our intentions are at the moment exploratory: we have chosen an ion on the basis of current laser technology, rather than on the basis of which ion which will give the best performance for the quantum computer. Other species of ion may well give better performance: In particular Beryllium ions do have the potential for a significantly lower error rate due to spontaneous emission, although it is also true that lighter ions may be more susceptible to heating. Other variations, such as the use of Raman transitions in place of single laser transitions, or the use of standing wave lasers need to be investigated. Our choice of Calcium will allow us to explore these issues. Furthermore, calculations suggest that it should be possible to trap 20 or more Calcium ions in a linear configuration and manipulate their quantum states by lasers on short enough time scales that many quantum logic operations may be performed before coherence is lost. Only by experiment can the theoretical estimates of performance be confirmed [36,37]. Until all of the sources of experimental error in real devices are thoroughly investigated, it will be impossible to determine what ion and addressing scheme enables one to build the best quantum computer or, indeed, whether it is possible to build a useful quantum computer with cold trap ions at all.

Acknowledgments

This research was funded by the National Security Agency.

References

1. R. J. Hughes et al. *Contemp. Phys.* **36** (1995) 149-163.
2. P. A. Benioff, *Int. J. Theor. Phys.* **21** (1982) 177-201.
3. R. P. Feynman, *Foundations of Physics* **16** (1986) 507-531.
4. D. Deutsch, *Proc. R. Soc. Lond.* **A 425** (1989) 73-90.
5. P. W. Shor, *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, S. Goldwasser ed., IEEE Computer Society Press, Los Alamitos CA, 1994.
6. A. Ekert and R. Jozsa, *Rev. Mod. Phys.* **68** (1996) 733-753.
7. R. J. Hughes, "Crptography, Quantum Computation and Trapped Ions", submitted to *Phil. Trans. Roy. Soc. (London)*, 1997; quant-ph/9712054.
8. J. I. Cirac and P. Zoller, *Phys. Rev. Lett.* **74**, (1995) 4094-4097.
9. C. Monroe et al., *Phys. Rev. Lett.* **75** (1995) 4714-4717.
10. E. Knill, R. Laflamme and W. Zurek, "Accuracy threshold for quantum computation"; submitted to *Science* (1997).
11. J. Preskill, "Reliable quantum computers", preprint (1997), quant-ph/9705031.
12. A. M. Steane, *Applied Physics B* **64** (1997) 623-642.

13. D. F. V. James, "Quantum dynamics of cold trapped ions, with application to quantum computation", Applied Physics B, in the press (1998); quant-ph/9702053.
14. D. J. Wineland et al., "Experimental issues in coherent quantum-state manipulation of trapped atomic ions", to be submitted to Rev. Mod. Phys. (1997).
15. R. J. Hughes, et al., "The Los Alamos Trapped Ion Quantum Computer Experiment", Fortschritte der Physik, in the press (1998); quant-ph/9708050.
16. L. K. Grover *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing*, ACM Press, New York, 1996 p.212
17. B. M. Terhal and J. A. Smolin, " Superfast quantum algorithms for coin weighing and binary search problems", preprint (1997) quant-ph/9705041.
18. D. Boneh and R. Lipton, "Quantum cryptanalysis of hidden linear functions," Proc. CRYPTO'95 (Springer, New York, 1995)
19. A. Kitaev, "Quantum measurements and the Abelian stabilizer problem," preprint (1995) quant-ph 9511026.
20. T. Pellizzari et al., Phys. Rev. Lett. **75** (1995) 3788-3791.
21. Q. A. Turchette et al., Phys. Rev. Lett. **75** (1995) 4710-4713.
22. D. G. Cory, A. F. Fahmy and T. F. Havel, Proc. Natl. Acad. Sci. USA **94** (1997) 1634-1639.
23. J. R. Friedman et al., Phys. Rev. Lett. **76** (1996) 3830-3833.
24. V. Privman, I. D. Vagner and G. Kventsel, "Quantum computation in quantum-Hall systems", preprint (1997), quant-ph/9707017.
25. M. F. Bocko, A. M. Herr and M. J. Feldman, "Prospects for quantum coherent computation using superconducting electronics", preprint (1997).
26. D. Loss and D. P. DiVincenzo, "Quantum computation with quantum dots" preprint (1997), cond-mat/9701055.
27. C. Monroe et al. Phys. Rev. **A 55** (1997) R2489.
28. A. M. Steane, private communication, 1996.
29. R. Blatt, private communication, 1997.
30. W. Paul and H. Steinwedel, Z. Naturforsch. **A 8** (1953) 448.
31. M. G. Raizen et al., Phys. Rev. **A 45** (1992) 6493-6501.
32. P. K. Ghosh, *Ion Traps* Clarendon Press, 1995.
33. F. Diedrich et al., Phys. Rev. Let. **62** (1989) 403-407.
34. S. Stenholm, Rev. Mod. Phys. **58** (1986) 699-739.
35. D. J. Wineland and W. M. Itano, Phys. Rev. **A 20** (1979) 1521-1540.
36. R. J. Hughes, D. F. V. James, E. H. Knill, R. Laflamme and A. G. Petschek, Phys. Rev. Lett. **77** (1996) 3240-3243.
37. D. F. V. James, R. J. Hughes, E. H. Knill, R. Laflamme and A. G. Petschek, *Photonic Quantum Computing*, S. P. Hotaling, A. R. Pirich eds, *Proceedings of SPIE* **3076** (1997) 42-50.

Arrays of Elliptical Ion Traps for Parallel Quantum Computing

Ralph G. DeVoe

IBM Almaden Research Center, 650 Harry Rd., San Jose CA 95120

Abstract. A new microscopic trap, the elliptical ion trap, can confine more than 10^3 ions with minimal micromotion, due to previously unrecognized properties of one-dimensional Coulomb crystals. Arrays of elliptical ion traps provide a physical model of a parallel quantum computer, consisting of thousands of microscopic Cirac-Zoller processors, each containing approximately 100 ions, maintained in entangled states by quantum optical interconnects. Physical properties of a trap array containing 4×10^6 qubits are derived.

Cirac and Zoller have recently proposed a model of a quantum computer [1,2] based on a linear ion trap. A row of ions are confined along the z-axis of the trap and laser cooled to the ground vibrational state. The lowest center-of-mass (c.m.) vibrational phonon of the ion array couples to the internal atomic states (the quantum bits or qubits) via Doppler shifted laser excitation. This phonon is used as a quantum communication channel to transmit superposition states between the atoms and carry out quantum logic. The Cirac-Zoller model contains all the interactions required for an evaluation of Shor's algorithm [3] and has been studied in sufficient detail [4] that it is often treated as a paradigm for the practicality of quantum computing in general. Several groups are constructing small scale tests of this model.

The two currently known quantum algorithms, Shor's factorizing algorithm and Grover's quantum database search [5], both require large numbers of qubits to gain an advantage over classical methods. Factorization of a 400 bit number has been estimated [4] to require > 2000 ions in the Cirac-Zoller model without quantum error correction (QEC) and from 10^4 to 10^6 ions with QEC [6]. Similarly, Grover's algorithm searches N unordered memory elements in only \sqrt{N} steps, and cannot compete with classical methods unless N is very large. However, experience with current linear ion traps is limited to $\approx 10^2$ ions [7] rather than the 10^4 to 10^6 needed.

The paper introduces a new trap, the elliptical ion trap, which overcomes some of the limitations of the linear ion trap. It has three important features. First, it acts like a linear trap for more than 10^3 ions, so that the c.m. phonon can be used for quantum logic as before. Second, the trap requires only a single flat electrode of microscopic dimensions so that trap arrays can be microfabricated at a density of $\geq 10^2$ traps/cm². Third, the size and geometry of the trap are consistent with previously proposed quantum optical interconnects [8,9,10,11],

since the trap is small enough to be enclosed by an optical cavity which can satisfy the "strong-coupling" condition of cavity quantum electrodynamics. The intention is that both the trap and the optical interconnects be microfabricated in large arrays, in analogy to conventional microprocessors. This model pictures a quantum computer as consisting of thousands of microscopic Cirac-Zoller processors, each containing approximately 100 ions, maintained in entangled states by quantum optical interconnects. Moreover, the trap array supports a parallel architecture in which the n bits of a large quantum word reside in n separate traps, as discussed in more detail below.

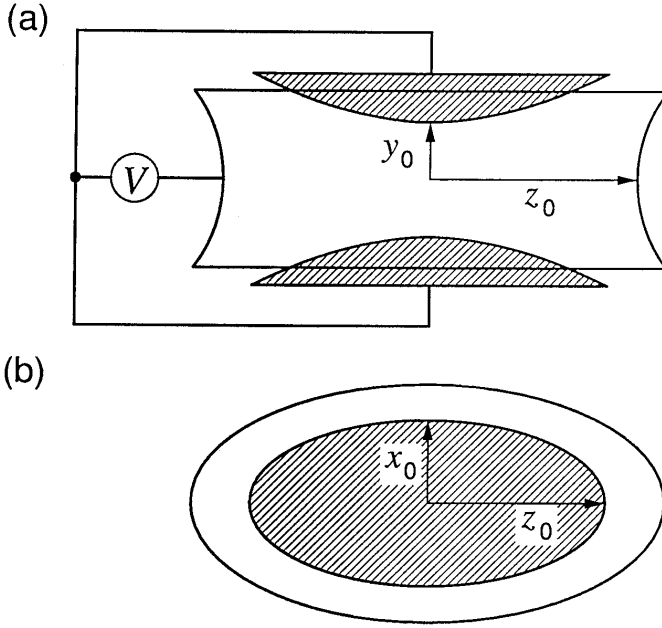


Fig. 1. An ideal elliptical ion trap, consisting of three elliptical hyperboloids. It generates the potential of Eq. 1 throughout the interior volume, since the electrodes lie on equipotential surfaces. The ellipticity $\epsilon = z_0/x_0$. In practice the simpler electrode of Fig. 4 suffices.

The elliptical trap, shown in Fig. 1, is a generalized ion trap which includes both the Paul trap [12] and the linear trap [13] as special cases. It is defined by the potential

$$V(x, y, z) = (V_{dc} + V_0 \cos(\Omega t)) \left(\frac{x^2}{x_0^2} - \frac{y^2}{y_0^2} + \frac{z^2}{z_0^2} \right). \quad (1)$$

The axes are named in agreement with the linear trap convention rather than the Paul trap convention. This is the most general quadratic potential satisfying $\nabla^2 V = 0$, which requires $1/x_0^2 + 1/z_0^2 = 1/y_0^2$. Such a potential can be created by hyperboloidal electrodes which form ellipses in the $x-z$ plane, where x_0 and z_0 are the minor and major axes of the ellipse at $y = 0$, while y_0 is the distance to the endcap electrode at $x = z = 0$. Solution of Newton's law for the above potential leads to three Mathieu equations [14,15] for the trapped ion's coordinates $u = x, y, z$

$$\frac{d^2 u}{d\tau^2} + \alpha_u [a_u - 2q_u \cos(2\tau)] u = 0 \quad (2)$$

where $\alpha_u = 1$ for $u = x, z$ and $\alpha_u = -1$ for $u = y$, the normalized time $\tau = \Omega t/2$, the Mathieu parameter (a dimensionless trap strength) is

$$q_u = -\frac{4eV_0}{m\Omega^2 u_0^2}, \quad (3)$$

m is the mass of the particle, $\Omega/2\pi$ the rf drive frequency, and $u_0 = x_0, y_0, z_0$. An identical equation holds for a_u with V_0 replaced by $2V_{dc}$. $\nabla^2 V = 0$ implies the relation $q_x + q_z = q_y$ and $a_x + a_z = a_y$. Define the ellipticity $\epsilon = z_0/x_0$ to be ratio of the major to minor axes of the ellipse which yields $q_z = q_y/(\epsilon^2 + 1)$ and $q_x = q_y\epsilon^2/(\epsilon^2 + 1)$. In the low q limit ($q < .2$) where most traps are operated [15] the trap secular frequencies $\omega_u \rightarrow q_u \Omega/2\sqrt{2}$ and therefore obey

$$\omega_z = \omega_y \frac{1}{\epsilon^2 + 1} \quad (4)$$

and

$$\omega_x = \omega_y \frac{\epsilon^2}{\epsilon^2 + 1}. \quad (5)$$

The Paul trap is a special case of $\epsilon = 1$ where $\omega_x = \omega_z = \omega_y/2$, while the linear trap is a limiting case of $\epsilon \rightarrow \infty$ where $\omega_z \rightarrow 0$ and $\omega_x \rightarrow \omega_y$. Although it is well-known that the linear trap is a limiting case of a stretched circular trap [13], and although the general form of Eq. 1 was discussed by Paul [14] many years ago, the regime $1 < \epsilon < \infty$ has not previously been considered in detail.

To play the role of a linear trap in the Cirac-Zoller model, an elliptical trap must first confine the ions on the z -axis so that the c.m. phonon can act as a quantum communications channel and second have sufficiently small micromotion so that the resulting Doppler shifts do not decouple the ions from the lasers. Schiffer [16,17], Dubin [18,19], and others have shown through analytic and numerical studies of ion crystals in asymmetric harmonic potentials that the ions will stay on the z -axis providing

$$d > d_0 \equiv \left(\frac{\alpha e^2}{m\omega_x^2}\right)^{1/3} \quad (6)$$

where d is the ion-ion separation at the center of the trap and $\alpha = 7\zeta(3)/2 = 4.207$, where ζ is the Riemann ζ -function. For $d < d_0$ the ions undergo what

is called the "zig-zag transition", where alternate ions move off the z-axis in opposite directions. Note that Eq. 6 does not depend directly on the axial trap strength ω_z , but only on the transverse trap strength ω_x . No exact analytic expression for d exists for all N but several accurate approximations have been derived. Equating Steane's [2] result $d = 2.0s_0N^{-0.57}$ to d_0 in Eq. 6, where the scale length $s_0^3 = e^2/m\omega_z^2$, yields the relation

$$N(\epsilon) = 1.44\epsilon^{2.34} \quad (7)$$

which determines the capacity of a trap of given ellipticity ϵ to hold a number $N(\epsilon)$ of ions on the z-axis. See Fig. 2.

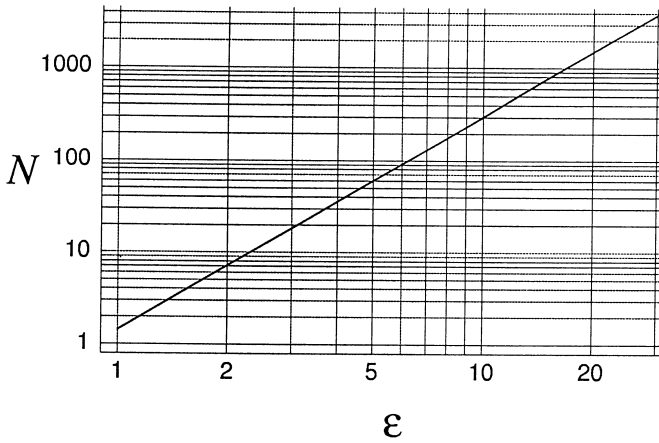


Fig. 2. A plot of Eq. 7, showing $N(\epsilon)$, the maximum number of ions which can be confined on the z-axis before the zig-zag transition, as a function of the ellipticity ϵ . Note that $N(\epsilon)$ is independent of all other trap parameters.

For $N > N(\epsilon)$ the ions assume the zig-zag structure while for $N < N(\epsilon)$ the ions are farther apart than the minimum value d_0 . Note that Eq. 7 is independent of all trap parameters except the ellipticity.

The second requirement is that the elliptical trap confine large crystals with minimal micromotion, that is, less than $\lambda/2\pi$, where λ is the wavelength of the laser light. It has been assumed previously that only linear ion traps, which have no r.f. field along the z-axis, can provide small enough micromotion for quantum computation. However, a detailed calculation of micromotion amplitudes shows that a properly designed elliptical trap can limit micromotion oscillations to a few tens of nm, even for large ion crystals containing more than 1000 ions and even though r.f. fields are used for all confinement. Micromotion [15,20] is the oscillatory motion of the ions at the rf drive frequency Ω and in the low q limit

is given by $\mu_z(t) = \bar{z}q_z \cos(\Omega t)/2$ where \bar{z} is the time-averaged (secular) position of the ion. The largest micromotion occurs at the ends where $\bar{z} = l_z/2$, where l_z is the total length of the ion chain. The theories of Dubin and Shiffer assume a static harmonic potential (the pseudopotential approximation) which yields accurate values of l_z but gives no information about micromotion. These results may be summarized by the approximation $l_z = d(N-1)^{1.053}$ which agrees within 10% with numerical results[18,21]. The micromotion may then be estimated using Steane's approximation for d yielding $\mu_z(N) = q_z s_0 (N-1)^{1.053}/2N^{.57}$ which approximates $q_z s_0 \sqrt{N}/2$ as $N \rightarrow \infty$. However, a much lower value of μ_z can be attained by matching the ellipticity to N . Consider a series of traps which are identical except for their ellipticity and assume that each trap contains $N(\epsilon)$ ions in accordance with Eq. 7. This minimizes micromotion for a given N by using the weakest axial confinement which will keep the ions on the z -axis. Using Eq. 4,5,6, and 7 yields

$$\mu_z(N) = \mu_z(2) \frac{1.985(N-1)^{1.053}}{N^{.569}(1+.733N^{.854})^{1/3}} \quad (8)$$

where $\mu_z(2)$ is the micromotion of 2 ions in a trap with $\epsilon = 1$. In the large N ($N \geq 20$) limit $\mu_z(N) \rightarrow 2.20N^{.20}\mu_z(2)$, that is, it depends approximately on the fifth root of N . This weak dependence is shown in Fig. 3 and permits trap parameters to be chosen so that micromotion is of little significance throughout the range $N=2$ to 1000, as shown in an example below.

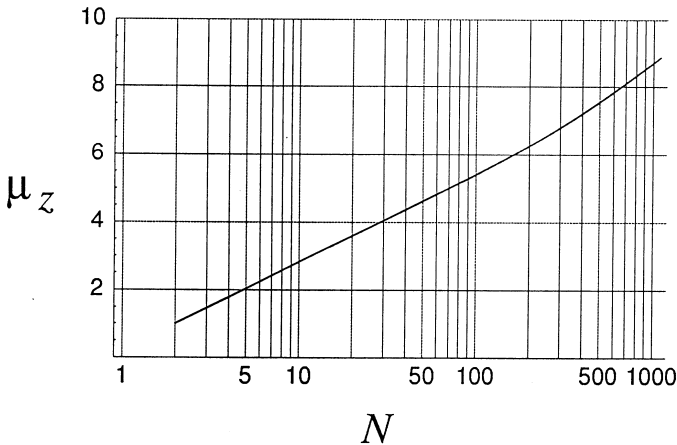


Fig. 3. The maximum micromotion amplitude μ_z in a linear crystal of N ions, Eq. 9. Note the weak dependence on N , which approximates $N^{1/5}$ at large N , due to previously unrecognized properties of 1-dimensional Coulomb crystals.

This result is unexpected since in three dimensions the micromotion amplitude increases approximately as $N^{1/3}$, due to the constant ion density [13]. Why does one-dimensional confinement result in a weaker power law of $N^{1/5}$? The answer is that in one dimension the density increases with N , due to more effective cancellation of the Coulomb repulsion of neighboring ions. A related result is that traps filled according to Eq. 7 all have the same value of $d = d_0$, despite the reduction of the axial trap force with increasing ϵ .

An elliptical trap can be constructed from single flat electrode, for example, an elliptical aperture in a conducting surface as shown in Fig. 4, which is straightforward to microfabricate in arrays by photolithography. This is in contrast to a

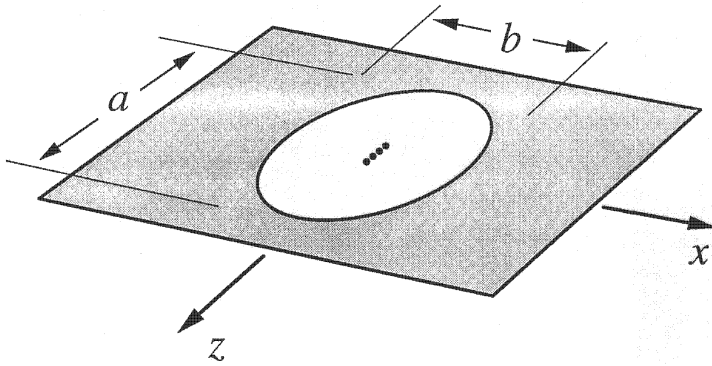


Fig. 4. An "aperture trap" consisting of an elliptical hole in a conducting surface which generates the potential of Eq. 1 near the z -axis. For the case considered in the text trap arrays can be generated a density of ≥ 100 traps/cm².

linear trap, which consists of a three dimensional array of four segmented rods [7]. The elliptical aperture generates a potential which reduces to Eq. 1 on the z -axis, where the ion crystal resides. The theory of such "aperture traps" has been discussed in detail [22] for the circular case and they have already been used in quantum optics [23] and quantum logic experiments [24]. In the circular case the trap parameters can be derived as coefficients in the Legendre expansion of the potential [22], but for $\epsilon > 1$ the corresponding power series contains ellipsoidal harmonics, which are rarely used. A three-dimensional numerical routine has therefore been used to estimate the three relevant parameters of the potential: the efficiency, the ellipticity, and the anharmonic terms. Potentials have been computed for apertures with ellipticities of 1,3,5, and 10, corresponding to $N(\epsilon)$ of 2, 18, 60, and 300 in Eq. 7. The efficiency is defined as $E = V_0/V_t$, where V_t is the potential applied to the conducting surface. The routine yields $.05 < E < .10$ depending on the thickness of the electrodes. The ellipticity of the potential is found to be equal to the ellipticity of the aperture within the accuracy of 20%

and the anharmonic terms were found to be comparable to the circular case[22]. Specifically for a 100 μm by 300 μm aperture, the z-axis fourth order (anharmonic) term was less than 1% of the second order term along a central 100 μm region of the z-axis. It is not apparent from numerical work whether an elliptical aperture produces the most accurate approximation to Eq. 1, since simpler apertures such as rectangles produce similar fields.

The practicality of elliptical traps can be demonstrated by considering an example based on recent experiments. Consider first an $\epsilon = 1$ trap of 100 μm radius, similar to ref. [23,24]. To separately excite each ion for quantum logic the ion-ion spacing d_0 must be greater than λ . Choosing $d_0 = 4 \mu\text{m}$ yields an axial oscillation frequency $\omega_z/2\pi \approx 1$ MHz at an applied potential $V_t=1000$ V rms (efficiency $E=.08$) for the ion Ba 137, currently under study. The Mathieu parameter $q_z = 2\sqrt{2}\omega_z/\Omega = 0.01$ for $\Omega/2\pi = 282$ MHz. This yields a micromotion $\mu_z = 10$ nm for a two ion crystal. Now assume that for quantum computing we wish to confine 64 ions. Eq. 7 yields an optimum ellipticity $\epsilon = 5.2$ which determines $q_z = 7.1 \times 10^{-4}$. The crystal length $l_z/2 \approx 32d_0 = 128$ microns, which gives $\mu_z = 45$ nm at the ends of the crystal and proportionately less near the center. The Doppler sidebands will therefore have an intensity of $J_1^2(\beta) \leq .06$, where $\beta = 2\pi\mu_z/\lambda$. A larger μ_z has been used in a recent experiment which observed interference and superradiance of two trapped ions[23]. The effect of such small micromotion amplitudes on quantum computing is therefore expected to be minimal.

The above 64-qubit trap occupies an area of $< 2 \times 10^{-3} \text{ cm}^2$ and may be placed in an array at a density of ≈ 100 traps/ cm^2 , yielding a bit density of 6400 qubits/ cm^2 . A 30 cm diameter disk could therefore support 64,000 traps holding 4×10^6 qubits. Such a disk driven at 1000 V rms at 282 MHz will dissipate only 3 watts, assuming a $Q=5000$. The laser power required to excite quantum logic in large arrays is also nominal, since recent experiments have saturated single ions with $\approx 10^{-7}$ watts of resonant light.

Quantum optical interconnects are essential for transferring quantum information between separate traps in an array. Recent theoretical work has proposed a variety of techniques [8,9,10,11] which transfer a quantum state between two distant atoms. These methods typically assume the "strong coupling" condition of cavity quantum electrodynamics[25], which requires that the cavity mode volume V_m obey $V_m < V_r$, where $V_r = \sigma_0 c/\Gamma$ is the atomic radiative volume, $\sigma_0 = 3\lambda^2/2\pi$, Γ is the atom's spontaneous emission decay rate, and c is the speed of light. Since $V_r \approx (100\mu\text{m})^3$ for ions of interest, cavity lengths must be less than one mm. This is consistent with the elliptical aperture traps of Fig. 4 since their two-dimensional structure permits dielectric mirrors to be brought within several hundred microns of the trap center. It would of course be desirable to also microfabricate the cavity mirrors in arrays, and one can imagine a three-disk "sandwich" structure in which the central disk contains the trap arrays and two outer integrated optical disks contain mirror arrays for the quantum optical interconnects.

This work was motivated by the issue of parallel architecture in quantum computation, specifically whether the very large number of gate operations (10^{11}) required by the modular exponentiation [4] in Shor's algorithm can be attributed to the serial architecture of the Cirac-Zoller model. Note that it is important to distinguish between quantum parallelism (entanglement) as defined by Deutsch, which is common to all quantum computers, and architectural parallelism, which is the simultaneous use of multiple quantum channels. The Cirac-Zoller model is serial since it relies on a single quantum channel, the c.m. phonon, for all communications. Preskill et. al. evaluate the modular exponentiation as series of repeated multiplications, which are in turn formed from repeated additions, again formed from repeated single-bit operations. This choice is constrained by the serial architecture. Can parallel algorithms evaluate the modular exponentiation with far fewer gates, thereby making decoherence less of an issue? Parallel quantum logic can also increase the execution speed (a limitation of the Cirac-Zoller model), in the classical case typically by a factor of $n \log n$, and is needed as well for quantum error correction [6]. Arrays of elliptical ion traps provide both a model system for the analysis of such issues and an avenue for experimental work. An elliptical trap is under construction in our laboratory, for use with laser-cooled Ba 137 ions.

I have benefited from many helpful comments on the manuscript from C. Kurtsiefer and D. DiVincenzo.

References

1. J.I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).
2. A. Steane, Appl. Phys. **B64**, 623 (1997).
3. P.W. Shor, in Proceedings of the 35th Annual Symposium on Foundations of Computer Science, (IEEE Computer Society Press, New York, 1994), p. 124.
4. D. Beckman, A. N. Chari, S Devabhaktuni, and J. Preskill, Phys. Rev. **A54**, 1034 (1996).
5. L.K. Grover, Phys. Rev. Lett. **79**, 325 (1997).
6. J. Preskill, Report No. quant-ph/9705031
7. M.G. Raizen, J.M. Gilligan, J.C. Bergquist, W.M. Itano, and D.J. Wineland, Phys. Rev. **A45**, 6493 (1992).
8. T. Pellizzari, S.A. Gardiner, J.I. Cirac, and P. Zoller, Phys. Rev. Lett. **75**, 3788 (1995).
9. J.I. Cirac, P. Zoller, H.J. Kimble, and H. Mabuchi Phys. Rev. Lett. **78**, 3221 (1997).
10. S.J. van Enk, J.I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 4293 (1997).
11. T. Pellizzari, Report No. quant-ph/9707001
12. W. Paul, Rev. Mod. Phys. **62**, 531 (1990).
13. J. Prestage, J. Appl. Phys. **66**, 1013 (1989).
14. W. Paul and M. Raether, Zeit. f. Physik **140**, 262 (1955).
15. R.F. Wuerker, H. Shelton, and R.V. Langmuir, J. Appl. Phys. **30**, 342 (1959).
16. J.P. Schiffer, Phys. Rev. Lett. **70**, 818 (1993).
17. R.W. Hasse and J.P. Schiffer, Ann. Phys. **203**, 419 (1990).
18. D.H. Dubin, Phys. Rev. Lett. **71**, 2753 (1993).
19. D.H. Dubin, Phys. Rev. **E55**, 4017 (1997).

20. H.G. Dehmelt, Adv. Atomic and Molecular Physics, **3**, 53 (1967).
21. D. James, Report No. quant-ph/9702053
22. R.G. Brewer, R.G. DeVoe, and R. Kallenbach, Phys. Rev. **A46**, 6781 (1992).
23. R.G. DeVoe and R.G. Brewer, Phys. Rev. Lett. **76**, 2049 (1996).
24. C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano, and D.J. Wineland, Phys. Rev. Lett. **75**, 4714 (1997).
25. H.J. Kimble in "Cavity Quantum Electrodynamics", P. Berman ed. (Academic, San Diego, CA 1994), p. 203

Simulating the Effect of Decoherence and Inaccuracies on a Quantum Computer

Kevin M. Obenland and Alvin M. Despain

Information Sciences Institute
[obenland,despain]@isi.edu

Abstract. A Quantum Computer is a new type of computer which can solve problems such as factoring and database search very efficiently. The usefulness of a quantum computer is limited by the effect of two different types of errors, decoherence and inaccuracies. In this paper we show the results of simulations of a quantum computer which consider both decoherence and inaccuracies. We simulate circuits which factor the numbers 15, 21, 35, and 57 as well as circuits which use database search to solve the circuit satisfaction problem. Our simulations show that the error rate per gate is on the order of 10^{-6} for a trapped ion quantum computer whose noise is kept below $\pi/4096$ per gate and with a decoherence rate of 10^{-6} . This is an important bound because previous studies have shown that a quantum computer can factor more efficiently than a classical computer if the error rate is of order 10^{-6} .

Keywords: Quantum Simulation, Ion Trap, Factoring, Database Search

1 Introduction

A quantum computer consists of atomic particles which obey the laws of quantum mechanics [TuHo95] [Lloy95]. The complexity of a quantum system is exponential with respect to the number of particles. Performing computation using these quantum particles results in an exponential amount of calculation in a polynomial amount of space and time [Feyn85] [Deut85]. This quantum parallelism is only applicable in a limited domain. Prime factorization is one such problem which can make effective use of quantum parallelism [Shor94]. This is an important problem because the security of the RSA public-key cryptosystem relies on the fact that prime factorization is computationally difficult [RiSA78].

Errors limit the effectiveness of any physical realization of a quantum computer. A quantum computer is subject to two different types of errors, decoherence and inaccuracies. Decoherence occurs when a quantum computer interacts with the environment. This interaction destroys the quantum parallelism by turning a quantum calculation into a classical one. The other type of error, inaccuracies in the implementation of gate operations, accumulates over time and destroys the results of the calculation.

In this paper we show results of simulations of a quantum computer which is subject to both decoherence and inaccuracies. These simulations assume the

trapped ion model of a quantum computer proposed by Cirac and Zoller [CiZo95]. We study Shor's factorization algorithm by simulating circuits which factor the numbers 15, 21, 35, and 57 [Shor94]. We also simulate Grover's database search algorithm with a circuit which solves the circuit satisfaction problem [Gro96]. The rest of this section gives a brief overview of quantum computers.

1.1 Qubits and Quantum Superposition

The basic unit of storage in a Quantum Computer is the *qubit*. A qubit is like a classical bit in that it can be in two states, zero or one. The qubit differs from the classical bit in that, because of the properties of quantum mechanics, it can be in both these states simultaneously [FeLS65]. A qubit which contains both the zero and one values is said to be in the superposition of the zero and one states. The superposition state persists until we perform an external measurement. This measurement operation forces the state to one of the two values. Because the measurement determines without doubt the value of the qubit, we must describe the possible states which exist before the measurement in terms of their probability of occurrence. These qubit probabilities must always sum to one because they represent all possible values for the qubit.

The quantum simulator represents the qubits of the computer using a complex vector space. Each state in the vector represents one of the possible values for the qubits. The bit values of a state are encoded as the index of that state in the vector. The simulator represents each encoded bit string with a non zero amplitude in the state vector. The probability of each state is defined as the square of this complex amplitude [FeLS65]. For a register with M qubits, the simulator uses a vector space of dimension 2^M .

1.2 Quantum Transformations and Logic Gates

A quantum computation is a sequence of transformations performed on the qubits contained in quantum registers [Feyn85] [BaBe95]. A transformation takes an input quantum state and produces a modified output quantum state. Typically we define transformations at the gate level, i.e. transformations which perform logic functions. The simulator performs each transformation by multiplying the 2^M dimensional vector by a $2^M \times 2^M$ transformation matrix.

The basic gate used in quantum computation is the controlled-not, i.e. exclusive or gate. The controlled-not gate is a two bit operation between a control bit and a resultant bit. The operation of the gate leaves the control bit unchanged, but conditionally flips the resultant bit based on the value of the control bit.

1.3 The Ion Trap Quantum Computer

The ion trap quantum computer as proposed by Cirac and Zoller is one of the most promising schemes for the experimental realization of a quantum computer [CiZo95]. Several experiments have demonstrated simple quantum gates

[MoMe95] [WiMM96]. Laser pulses directed at the ions in the trap cause transformations to their internal state. A common phonon vibration mode is used to communicate between the ions in the trap. A controlled-not gate is a sequence of laser pulses. We use the ion trap quantum computer as the model for our quantum simulator.

Qubits in the Ion Trap Quantum Computer. Qubits are represented using the internal energy states of the ions in the trap. The ion trap represents a logic zero with the ground state of an ion, and a logic one with a higher energy state. The ion trap quantum computer also requires a third state which it uses to implement the controlled-not gate. In this paper we use a simplified model which, instead of using a third state for each qubit, uses a single third state which is shared amongst all the qubits. This simplified model reduces the simulation complexity exponentially without an appreciable loss of accuracy [ObDe97a].

Transformations in the Ion Trap. An operation in the ion trap quantum computer is a sequence of laser pulses. Each laser pulse is defined by one of the transformation matrices shown in equation 1. θ corresponds to the duration of the laser pulse and ϕ corresponds to the phase. A two bit controlled-not gate is a sequence of five laser pulses, two V and three U transformations. A single bit not gate can be implemented with three laser pulses and the three bit controlled-not gate requires seven laser pulses.

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \frac{\theta}{2} & -ie^{-i\phi} \sin \frac{\theta}{2} & 0 \\ 0 & -ie^{i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad V = \begin{bmatrix} \cos \frac{\theta}{2} & -ie^{-i\phi} \sin \frac{\theta}{2} \\ -ie^{i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix} \quad (1)$$

2 Simulating a Quantum Computer

Our quantum computer simulator simulates circuits at the gate level. The simulator implements one, two and three bit controlled-not gates as well as rotation gates. The simulator implements each gate as a sequence of laser pulses, and represents the entire vector space throughout the simulation.

2.1 Operational Errors and Decoherence

The simulator models inaccuracies by adding a small deviation to the two angles of rotation θ and ϕ . Each *operational* error angle is drawn from a gaussian distribution with a parametrized mean (μ) and standard deviation (σ). Errors with non zero μ , called *mean* errors, correspond to systematic calibration errors, and errors with non zero σ , referred to as *standard deviation* errors, correspond to noise in the laser apparatus.

Because the phonon mode is coupled to all the qubits in the computer, it is the largest source of decoherence [MoMe95]. For this reason we only model the phonon decoherence and not the decoherence of the individual qubits.

We model the decoherence of the phonon mode by performing an additional operation after each laser pulse. Equation 2 shows this transformation which has the effect of decaying the amplitude of the states in the phonon state. This decay transformation is based on the quantum jump approach [Carm93]. The decay parameter (*dec*) remains constant throughout the entire simulation.

$$\begin{matrix} |\psi\rangle > |0\rangle_p \\ |\psi\rangle > |1\rangle_p \end{matrix} \Rightarrow e^{-dec/2} \begin{matrix} |\psi\rangle > |0\rangle_p \\ |\psi\rangle > |1\rangle_p \end{matrix} \quad (2)$$

This method of modeling decoherence implicitly models spontaneous emission. Because the state is never renormalized, the total norm at each step represents the probability that the calculation survives up to that point without a spontaneous emission occurring. An alternative method for modeling decoherence is to renormalize the state at each step and then, based on a probability of emission, cause emissions at different points in the calculation. This method has the disadvantage that, because we cause emissions at random points in the calculation, we must run multiple simulations each with different initial random seeds to average out any bias caused by the random number generator. We have shown however that both methods for modeling decoherence give essentially the same results [ObDe97a].

Because the simulator applies the decoherence transformation once per laser pulse, the parameter *dec* has units of (decoherence/laser pulse). To convert these units to decoherence per unit time we must consider the switching time of the laser. The *dec* parameter is simply the switching speed divided by the decoherence time. Recent experiments show switching speeds of 20kHz for a controlled-not gate, i.e. four π pulses, and a decoherence rate of a few kHz [MoMe95]. This corresponds to a decoherence parameter value between 10^{-2} and 10^{-3} .

2.2 Quantum Circuits

Much of the current interest in quantum computation is due to the discovery of an efficient algorithm by Peter Shor to factor numbers [Shor94]. By putting the qubit register A in the superposition of all values and calculating the function $f(A) = X^A \bmod N$, a quantum computer calculates all the values of $f(A)$ simultaneously. Where N is the number to be factored, and X is a randomly selected number which is relatively prime to N . The quantum factoring circuit also contains operations to create the superposition state at the beginning of the circuit, and extract the period at the end of the circuit. The circuit to calculate $f(A)$ can be performed in $O(L^3)$ time using repeated squaring [Desp96], i.e. a sequence of multiplications performed modulo N .

Grover's database search algorithm searches for a key, from a set of matching keys, in an unsorted database [Gro96]. The keys are defined by a function which can be evaluated in unit time. After evaluating this function a *diffusion*

transformation is performed which amplifies the probability in the states with matching keys. Grover shows that after performing $O(\sqrt{N})$ evaluation steps and diffusion transformations the probability of measuring a matching key is greater than $1/2$.

3 Simulation Results

In this section we study how decoherence and operational errors degrade the fidelity of the factoring and database search algorithms.

The fidelity, as defined by $fidelity = \| \langle \varphi | \psi \rangle \|^2$, measures how close a state with error in it is to the correct result. The fidelity is defined as the inner product between the simulation with errors (ψ) and the correct result (φ).

Table 1 shows all the benchmarks used in our simulation studies. To show the complexity of simulating these benchmarks we show the simulation time for each. This time assumes a single 300 MHz processor and the simulations include only operational error. All the factoring benchmarks were run using a parallel version of the simulator [ObDe98]. We have run simulations on as many as 256 processors, and the simulator achieves near linear speedup.

Table 1. Benchmarks used in simulation studies

Benchmark	Number of qubits	Number of laser pulses	Description	Simulation time (secs)
grover	13	1,838	Circuit SAT using the Grover database search algorithm	10
mult	16	8,854	One modulo multiply step from the factor-15 problem	282
factor15	18	70,793	Factor-15 problem using 3 qubits for A	10,465
factor21	24	69,884	Factor-21 problem using 6 qubits for A	272,276
factor35	27	99,387	Factor-35 problem using 6 qubits for A	3,083,520
factor57	27	97,939	Factor-57 problem using 6 qubits for A	3,067,853

Because of round off error, in the factoring algorithm, the choice of the number of bits to use in the A register affects the probability seen after performing the FFT [Shor94]. Shor suggest using $2L + 1$ bits for an L bit factorization, but for the numbers used in our studies we can use less. For the factor15 problem the period is a power of two, i.e. four, and therefore there is no round off error. For the numbers 21, 35 and 57, the probability given by the FFT does not increase for more than six bits. Also we are mainly concerned with observing the fidelity for these circuits, and the fidelity is always calculated before the FFT.

3.1 Operational Errors

In this section we consider operational errors without any decoherence. We first investigate the significance of errors in the angle ϕ , by varying the amount of error in the angles θ and ϕ separately. In all other simulations we vary the angles θ and ϕ together.

To average the random bias out of simulations with non zero σ we run multiple simulations each with different initial random seeds. To get a single simulation point we run at least four simulations, and in many cases we run more to establish upper and lower confidence intervals for the average of the ending fidelities [HiMo80]. For simulations which include errors in the angle θ we run enough simulations so that the average of the fidelities is within 0.02 of the actual mean of the distribution with an upper and lower confidence of 95%. There is more variability in the fidelity for simulations which consider only ϕ errors, so for these we obtain the 95% confidence intervals to within 0.03 of the mean.

Significance of Errors in the Angle ϕ . Fig. 1 shows how operational errors degrade the fidelity for the factor15 benchmark. In Fig. 1(a) we vary the mean and standard deviation and introduce both θ and ϕ errors. In Fig. 1(b) we only introduce θ errors. Comparing the two graphs shows that the combination of θ and ϕ errors produces a lower fidelity than θ errors alone.

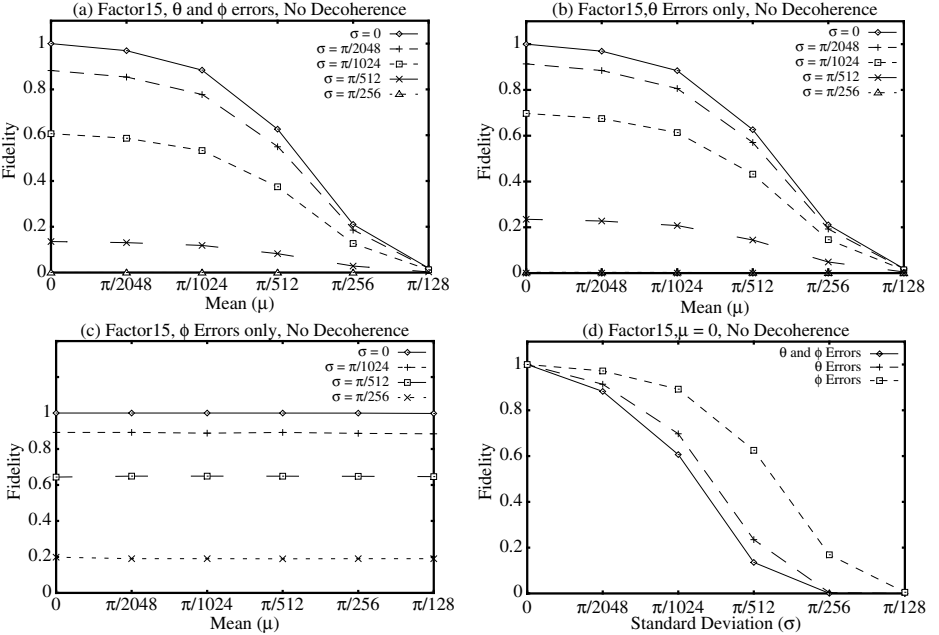


Fig. 1. Inaccuracies for factor15 (a) θ and ϕ errors (b) θ errors (c) ϕ errors (d) θ and ϕ errors, θ errors, ϕ errors

Fig. 1(c) shows the effect of only ϕ errors. As the graph shows noise degrades the fidelity, but adding a constant amount of error has no effect. Fixed magnitude ϕ errors have no effect because the laser transformations are always performed in pairs, and an error in the second transformation cancels an error in the first transformation. In Fig. 1(d) we compare the effect of θ and ϕ errors alone and their combined effect. The highest degradation occurs when considering both θ and ϕ errors, and θ errors produce a more significant effect than ϕ errors.

Operational Errors for the Grover Benchmark. Fig. 2 shows mean and standard deviation operational errors for the grover benchmark. The figures show the probability of finding a correct key for twelve iterations of the algorithm. When running the database search it is important to stop after the correct iteration, because the probability decreases if we run too many iterations [BoBr96]. For our case the highest probability occurs after the eighth iteration.

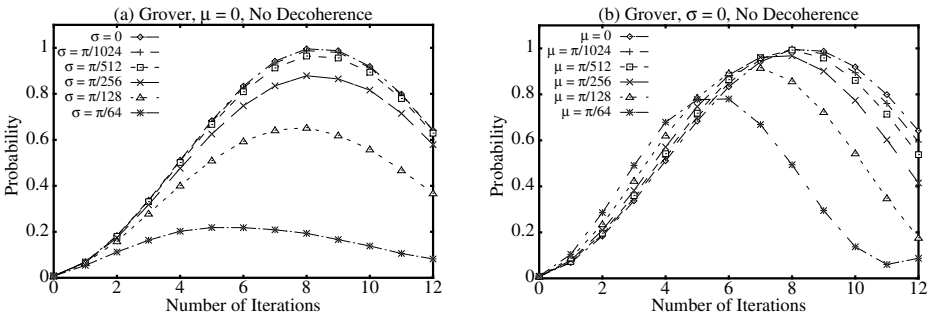


Fig. 2. Inaccuracies for the grover benchmark. (a) σ Errors. (b) μ errors

Fig. 2(a) shows that the peak probability is above 0.5 for σ errors as great as $\pi/128$, and for an error rate of $\pi/64$ the peak probability is 0.2. Also for σ errors the peak probability occurs after the same iteration for all errors less than $\pi/128$. However μ errors, as shown in Fig. 2(b), shift the peak so that it occurs at an earlier iteration. The peaks are higher for μ errors than they are for σ errors with the same level of error. However, since we can only perform a single measurement, the shift in the peak values causes a further reduction in the probability if we measure after the eighth iteration.

Fidelity at Intermediate Points for the Factoring Benchmarks. Fig. 3 shows the fidelity at intermediate points in the calculation for the factor21, and factor57 benchmarks.

Standard deviation errors produce a more significant effect than μ errors of the same magnitude. This is due to a cancellation effect for mean errors which is very similar to the cancellation effect exhibited by ϕ errors [ObDe97a]. This cancellation effect occurs because of the nature of reversible computation and

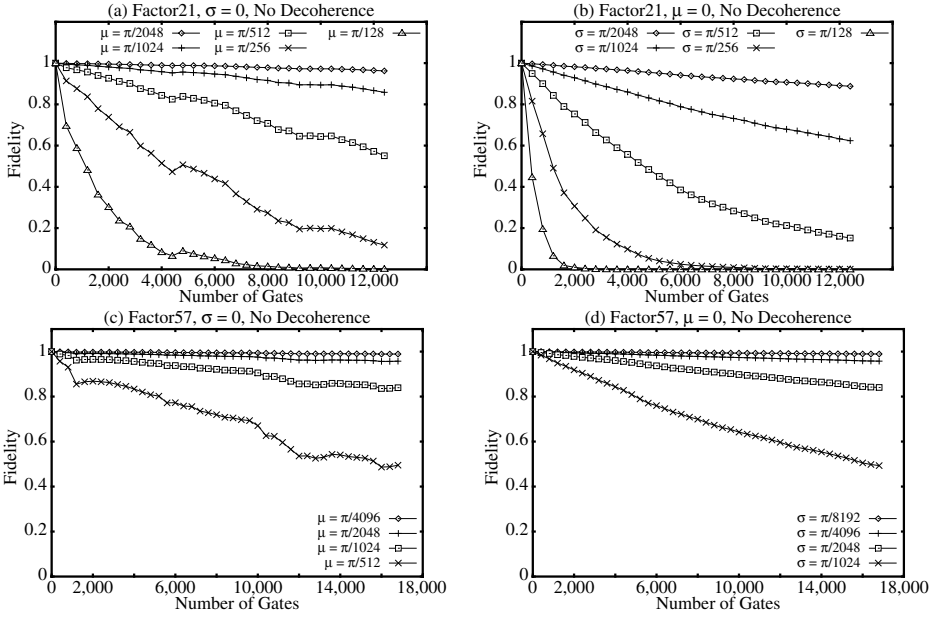


Fig. 3. Fidelity as a function of the number of gates for factor21 and factor57

causes the fidelity to go through periods where it increases. Because we perform operations in pairs, where the two operations are inverses of each other, an error in one operation may reverse the error from an earlier operation. Standard deviation errors also exhibit this effect but it is reduced because we average the results from multiple simulations.

3.2 Decoherence Errors

Fig. 4 shows how the fidelity of a computation decreases over time in the presence of decoherence. Fig. 4 shows the fidelity at intermediate points in the calculation for the four factoring benchmarks. For small amounts of decoherence, i.e. 10^{-6} or less, the fidelity is not adversely affected. A decoherence rate of 10^{-4} results in a steady decrease in the fidelity over the course of the computation, and for rates even higher the fidelity drops off very quickly. As the figure shows, decoherence has a similar effect on all of the benchmarks.

3.3 The Correlation Between Decoherence and Operational Errors

Both decoherence and operational error cause a degradation of the fidelity in a quantum computation. Decoherence degrades the fidelity through the decay of the phonon state, and operational errors result in the accumulation of amplitude in unwanted states. The combined effect of these two factors is a degradation

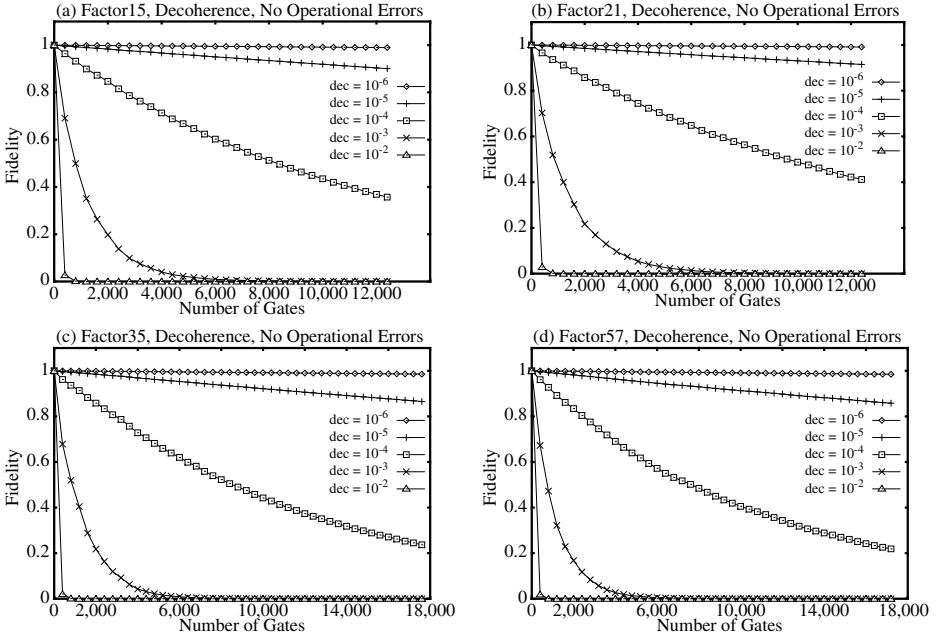


Fig. 4. Fidelity as a function of the number of gates for decoherence in the factor15, factor21, factor35 and factor57 benchmarks

which is worse than either factor considered alone. We can represent the combined effect as: $F_{dec,op} = F_{dec} \bullet F_{op} + \Omega(F_{dec}, F_{op})$ Where F_{dec} and F_{op} are the fidelities of simulations for decoherence and operational error considered separately, and $\Omega(F_{dec}, F_{op})$ is the correlation between the two types of error. As Table 2 shows the correlation is very low. We calculated the correlation by running simulations which considered decoherence and operational error together. For all the benchmarks the maximum correlation is at most 1.14×10^{-2} . This result means that we can simulate decoherence and operational errors separately, and combine the results to obtain their collective effect on a calculation.

The complexity of simulating decoherence alone is much lower than the complexity of simulating operational errors. This is because the decay transformation does not have any off diagonal terms, and therefore it does not introduce any new error states. The simulator only needs to represent enough states to represent the superposition state. Instead of using the index of a state to represent its bit value, the simulator now keeps an extra field for each state which holds the current value of the bit string for that state.

Using this new method for modeling decoherence, the simulator only needs to allocate $O(2^L)$ states to simulate the factorization of an L bit number. To represent all the qubits for this problem, the simulator would need to allocate $O(2^{4L})$ states. This reduces the memory requirements and simulation complexity

Table 2. Correlation (Ω) between decoherence and operational errors

Benchmark and Simulation Model	Maximum Ω	Average Ω
mult, $\mu=0$, $\sigma = \pi/1024 - \pi/64$	5.76×10^{-5}	3.63×10^{-6}
mult, $\sigma = 0$, $\mu = \pi/1024 - \pi/64$	9.26×10^{-3}	5.51×10^{-4}
factor15, $\sigma = \pi/1024$, $\mu = 0$	4.15×10^{-3}	3.96×10^{-4}
factor15, $\sigma = 0$, $\mu = \pi/1024$	1.14×10^{-2}	8.76×10^{-4}
grover, $\mu=0$, $\sigma = \pi/1024 - \pi/128$	1.78×10^{-3}	1.02×10^{-4}
grover, $\sigma=0$, $\mu = \pi/1024 - \pi/128$	2.67×10^{-3}	2.36×10^{-4}

by a factor of $O(2^{3L})$. For example the simulation of the factor15 circuit requires only 1/4096 the amount of time as before.

3.4 The Error Rate per Gate

As our simulation results show, a modest amount of error destroys even a relatively small calculation. If quantum computers are to be useful, we must be able to perform calculations which are even larger than the ones considered here. These larger calculations will therefore require the use of quantum error correcting codes [Stea96]. Several recent studies have shown that, by using fault tolerant techniques, if the error of an individual gate is low enough we can perform a useful quantum calculation of indefinite length [Pres96] [KnLZ96]. This *accuracy threshold* is expressed in terms of the probability of error per gate. We can use the results of our simulation studies to show how this error probability relates to decoherence and inaccuracies.

Table 3 shows the error rate per gate considering decoherence and operational errors for the factor57 benchmark. Error rates for the other factoring benchmarks as well as error rates for the combination of decoherence and operational errors are given in [ObDe97b]. The error rate is calculated as $(1 - \text{Fidelity})/\text{Number of Gates}$. To get the error rate for a particular amount of error, we calculate the error rate after every tenth gate in a computation and take the average. A gate is either a one, two or three bit controlled-not gate or a single bit rotation. It takes five laser pulses on average to implement each gate.

To perform a computation of arbitrary length the error rate must be about 10^{-5} for one and two bit gates and 10^{-3} for three bit gates [Pres96]. An error rate of 10^{-5} corresponds to operational errors with $\sigma = \pi/2048$ and decoherence of 10^{-5} . Table 3 shows that we can tolerate an even higher level of constant magnitude errors. The error rate is 10^{-5} for operational errors with $\mu = \pi/1024$.

To perform a quantum factorization which is more efficient than a classical one the error threshold is even tighter, roughly 10^{-6} . This corresponds to operational errors with $\sigma < \pi/4096$ and a decoherence rate of 10^{-6} .

Using the error rate of our factoring circuits to predict the error rate for error correction circuits assumes that these two types of circuits behave in a

Table 3. Average error rate per gate for the factor57 benchmark

Decoherence		Operational Errors, $\mu = 0$		Operational Errors, $\sigma = 0$	
dec	Error Rate	σ	Error Rate	μ	Error Rate
10^{-7}	9.1×10^{-8}	$\pi/8192$	6.6×10^{-7}	$\pi/4096$	7.9×10^{-7}
10^{-6}	9.1×10^{-7}	$\pi/4096$	2.6×10^{-6}	$\pi/2048$	3.1×10^{-6}
10^{-5}	8.8×10^{-6}	$\pi/2048$	1.0×10^{-5}	$\pi/1024$	1.2×10^{-5}
10^{-4}	6.5×10^{-5}	$\pi/1024$	3.6×10^{-5}	$\pi/512$	4.4×10^{-5}

similar manner. These circuits are similar because they are both built from the same types of elementary gates, i.e. controlled-not gates. Also the error rate, for a given amount of error, is very similar for all the factoring circuits that we considered. Lastly effects that we have seen such as error cancellation are a by-product of the fact that quantum circuits and gates are implemented in a reversible fashion. Because of the nature quantum mechanics quantum circuits will always be implemented in this way.

4 Conclusion

Quantum computation is a new type of computation which can achieve exponential parallelism. The feasibility of a quantum computer is threatened by two types of errors, decoherence and inaccuracies. In this paper we performed simulations of a quantum computer running Shor’s factoring algorithm and Grover’s database search algorithm to access the feasibility of implementing a quantum computer.

Our simulations show that random inaccuracies (noise) are more significant than fixed magnitude inaccuracies for the ion trap quantum computer. Also errors in the duration of the laser pulse are more significant than errors in the phase of the laser. For the problems considered in this paper we show that noise or constant magnitude inaccuracies of magnitude $\pi/512$ or greater cause a significant impact on the fidelity of the calculation.

Our simulations also show that a quantum computation can tolerate a decoherence rate as high as 10^{-5} . For the ion trap computer this corresponds roughly to a decoherence lifetime of 50 milliseconds and a switching speed of 500 nanoseconds. We also show that inaccuracies and decoherence are uncorrelated and can be simulated separately.

Our simulations relate the physical quantities of inaccuracies and decoherence to the probability of error per gate. An error rate per gate on the order of 10^{-6} corresponds to inaccuracies of less than $\pi/4096$ per laser operation and a decoherence rate of 10^{-6} . A quantum computer with this error rate, if it uses quantum error correcting codes, could factor a number more efficiently than a classical computer. This assumes that the quantum circuits used to implement

factoring with error correction codes behave in the same manner as the factoring circuits used in this paper.

Acknowledgments

The authors are members of the Quantum Information and Computation (QUIC) consortium. We wish to thank our QUIC colleagues: Jeff Kimble, John Preskill, Hideo Mabuchi and Dave Vernooy. This work is supported in part by ARPA under contract number DAAH04-96-1-0386.

References

- [BaBe95] A. Barenco et al. "Elementary Gates for Quantum Computation." *Phys. R. A* **52**, 3457-3467. 1995.
- [BoBr96] M. Boyer et al. "Tight bounds on quantum searching." *Proc. PhysComp96*.
- [Carm93] H.J. Carmichael. *An Open Systems Approach to Quantum Optics, Lecture notes in Physics*, (Springer, Berlin). 1993.
- [CiZo95] J.I. Cirac, and P. Zoller. "Quantum Computations with Cold Trapped Ions." *Phys. Rev. Lett.* **74**, Number 20. May 15, 1995.
- [Desp96] A. Despain. "Quantum Networks" in *Quantum Computing*. JASON Report **JSR-95-115**. pp 49-81. The MITRE Corp. 1996.
- [Deut85] D. Deutsch. "Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer." *Proc. R. Soc. Lond. A* **400**, pp 97-117. 1985.
- [FeLS65] R. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics III*. Addison-Wesley Publishing Company. 1965.
- [Feyn85] R. Feynman. "Quantum Mechanical Computers." *Found. of Phys.*, **16**, 1985.
- [Gro96] L. Grover. "A Fast Quantum Mechanical Algorithm for Database Search." *Proc., STOC* 1996.
- [HiMo80] W. W. Hines, D. C. Montgomery. *Probability and Statistics in Engineering and Management Science*. John Wiley & Sons, Inc. 1980.
- [KnLZ96] E. Knill et al. "Accuracy Threshold for Quantum Computation." 1996.
- [Lloy95] S. Lloyd. "Quantum-Mechanical Computers." *Scientific American*. Vol. **273**, No. 4, pp 140-145. October 1995.
- [MoMe95] C. Monroe et al. "Demonstration of a Universal Quantum Logic Gate." *Phys. Rev. Lett.* **75**, 4714. Dec. 1995.
- [ObDe97a] K. Obenland, and A. Despain. "Models to Reduce the Complexity of Simulating a Quantum Computer." *ISI Tech. Report*. November 1997.
- [ObDe97b] K. Obenland, and A. Despain. "Simulating the Effect of Decoherence and Inaccuracies on a Quantum Computer." *ISI Tech. Report*. December 1997.
- [ObDe98] K. Obenland, and A. Despain. "A Parallel Quantum Computer Simulator." *High Performance Computing 1998*.
- [Pres96] J. Preskill. "Reliable Quantum Computers." *Proc. R. Soc. Lond. A*. 1996.
- [RiSA78] R.L. Rivest et al. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems." *Comm. ACM*. **21**, p 120. 1978.
- [Shor94] P. Shor. "Algorithms for Quantum Computation: Discrete Logarithms and Factoring." *Proc., 35th Annual FOCS* pp. 124-134. November 1994.
- [Stea96] A.M. Steane. "Multiple Particle Interference and Quantum Error Correction." *Proc. R. Soc. Lond. A* **452**, 2551. 1996.

- [TuHo95] Q.A. Turchette et al. "Measurement of Conditional Phase Shifts for Quantum Logic." *Phys. Rev. Lett.* **75**, 4710. Dec. 1995.
- [WiMM96] D. Wineland et al. "Quantum State manipulation of Trapped Atomic Ions." *Proc. R. Soc. Lond. A*. December 1996.

Implementation of Quantum Controlled-NOT Gates Using Asymmetric Semiconductor Quantum Dots

Alexander A. Balandin and Kang L. Wang

Device Research Laboratory
Electrical Engineering Department
University of California, Los Angeles CA 90095, USA,
alexnb@ee.ucla.edu

Abstract. We propose an implementation of a quantum controlled-NOT gate on the basis of dipole-dipole interacting *asymmetric* quantum dots. Our implementation does not require application of an external electric field as the one proposed earlier [Barenco *et. al*, Phys. Rev. Lett., **74**, 4083 (1995)]. Results of our numerical simulations show that owing to the dot asymmetry, the coupling constant of the dipole-dipole interaction can be made as large as $\hbar\omega_d \approx 50$ meV while keeping the probability of the spontaneous emission low. This provides conditions for resolving different entangled quantum states experimentally.

Keywords: quantum gate, asymmetric quantum dot, dipole interaction

1 Introduction

Development of quantum computation schemes has led to many proposals of potentially realizable quantum computers based on various physical systems. Among the most known are quantum computers that utilize laser trapped atoms [1], nuclear magnetic resonance (NMR) systems [2], all-optical logic gates [3], and semiconductor nanostructures [4]. Successful experimental demonstrations of one and two qubit computers were reported for laser trapped atom systems and NMR systems [5]. The progress in experimental development of quantum logic gates on the basis of semiconductor nanostructures has been far short of these systems. This is primarily due to the difficulties in fabrication of high quality quantum dot arrays and the decoherence problem [6] which is inherently more severe for solid state systems. Meanwhile, if these problems are overcome, the quantum computer based on semiconductor quantum dots offers an attractive alternative to other physical implementations due to its compactness, robustness, and the larger number of qubits which can be realized [4], [7], [8]. To add to these merits, the solid state implementation of a quantum computer avoids dealing with a statistical mixture of pure quantum states like in NMR realizations [2], or pitfalls related to manipulation of single ions in a trap. Another important advantage of the quantum dot implementation is that further development can make use of appropriate modification of well-developed conventional silicon based technology.

In this paper we propose a solid state implementation of a quantum controlled-NOT gate based on coupled asymmetrical quantum dots. These dots have asymmetric potential profiles along one of the dots' axis such as the step quantum dot in Fig. 1. Other potential profiles which can be used for our structure were discussed in a different context in Refs. [11], [12]. The potential well asymmetry has the effect similar to that of applying an electric field but allows for more degrees of freedom. This is a major advantage for the quantum logic gates. The rest of the paper is organized as follows. In section 2 we present an example structure and the operational principle of the proposed solid-state implementation of quantum controlled-NOT gate; section 3 shows electron charge density distributions in asymmetric dots obtained by numerical simulation. Discussion and comparison of the performance of the proposed gate with other implementations are given in section 4. We present our conclusion in section 5.

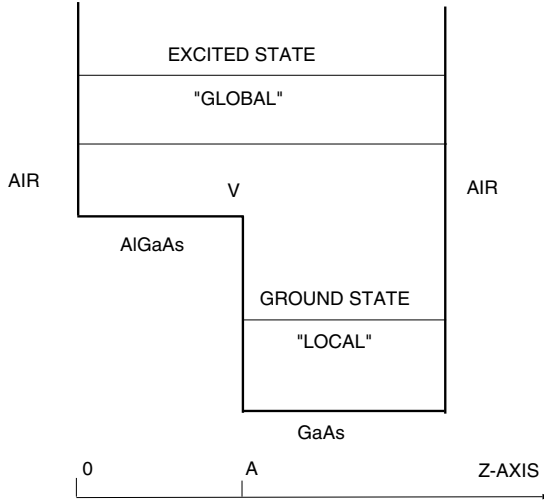


Fig. 1. Potential profile of the quantum dot along the z -axis. The step height of the potential profile and its position can be tuned by a proper choice of materials and growth conditions.

2 Structure of the Gate

We consider two single-electron quantum dots separated by a distance R . The ground state and one of the excited states j of each dot form computational basis states $|0\rangle$ and $|1\rangle$ in much the same way as it was suggested in Ref. [4]. The actual physical implementation of this gate may have more than one electron per dot. However, we restrict our consideration to one electron per dot for the

sake of simplicity. The first quantum dot with the resonant energy $\hbar\omega_c$ acts as the control qubit while the second dot with the resonant energy $\hbar\omega_t$ acts as the target qubit. The gate is driven optically by application of a proper train of π pulses [9]. The schematic of the potential profile of the dot is given in Fig. 1. The dots with such asymmetric potential profile can be grown by self-assembly on top of some pre-patterned substrate. Current state-of-the-art technology allows for fabrication of such dots with the feature size of 50Å- 300Å with a 50Å-200Å separation between two dots [10]. Since these dots are separated by air and not by the layer of some semiconductor, the electrical dipole-dipole coupling between them is stronger due to the lack of dielectric screening. It also allows us to assume hard-wall boundary conditions for the dots.

Due to the asymmetry of the potential profile of the quantum dot, the charge distribution in the ground state and the excited states of each dot is uneven in the absence of an external field. The charge distribution can also be tuned and shifted in any direction by a proper modification of the potential profile. The difference in the charge distributions between two states is particularly pronounced when one of the states is so-called "local state" and the other one is "global state" [11], [12]. We call the state "local" if its wave function is almost entirely localized in the vicinity of some smaller area inside the dot. In Fig. 1, only the ground state is the local state. We intentionally did not specify which of the excited states we use as the basis state $|1\rangle$ since for a particular potential profile and a dot size it may be better to utilize higher excited states rather than the first excited state. Owing to the dipole-dipole interaction, the resonant frequency for transitions between the excited and ground states in one dot depends on the electron state of the neighboring dot. The coupling constant of the dipole-dipole interaction can be written as [4]

$$\hbar\omega_d = -\frac{d_c d_t}{4\pi\epsilon_r\epsilon_o R^3}, \quad (1)$$

where \hbar is the Plank's constant, ϵ_o and ϵ_r are the vacuum and relative dielectric permittivities, respectively; d_c and d_t are dipole moments of the control and target dots, respectively. The resonant frequency for the target dot becomes $\omega_t \pm \omega_d$ depending on the state of the control dot ($|0\rangle$ or $|1\rangle$). As a result, a π pulse of frequency $\omega_t + \omega_d$ causes the transition $|0\rangle \rightarrow |1\rangle$ only provided that the control dot is in state $|1\rangle$.

In order to be able to determine the state of the system experimentally, the coupling constant $\hbar\omega_d$ has to be relatively large

$$\Delta\omega \equiv \omega_{t,c} - \omega_d \leq \Gamma, \quad (2)$$

where Γ is the emission (absorption) line broadening due to interaction with phonons and impurities. The probability of spontaneous transitions has also to be small and controlled via a proper choice of the basic computational states and structural engineering in order to be able to perform many computations.

3 Numerical Simulation

To determine the electron wave functions and resonant frequencies for each confined state, we applied the transfer matrix method [13]. The dipole due to the uneven charge distributions was then calculated as

$$d = \int_{-L/2}^{L/2} \Psi(z) z \Psi^*(z) dz, \quad (3)$$

where L is the dot's size along z -axis, and $\Psi(z)$ is the component of the electron wave function envelope. Material parameters used in simulations correspond to *GaAs*. The electron effective mass is assumed to be $0.067m_o$, where m_o is the free electron mass. The relative permittivity of *GaAs* is taken to be 12.9. The charge density distributions for the first four confined states together with corresponding confined energies and dipole moments are shown in Fig. 2 and Fig. 3 for two different dot sizes. In both cases, the step (point A in Fig. 1) is located at the middle of the dot. The height of the potential step approximately corresponds to the heterojunction band offset between *GaAs* and $Al_{0.2}Ga_{0.8}As$.

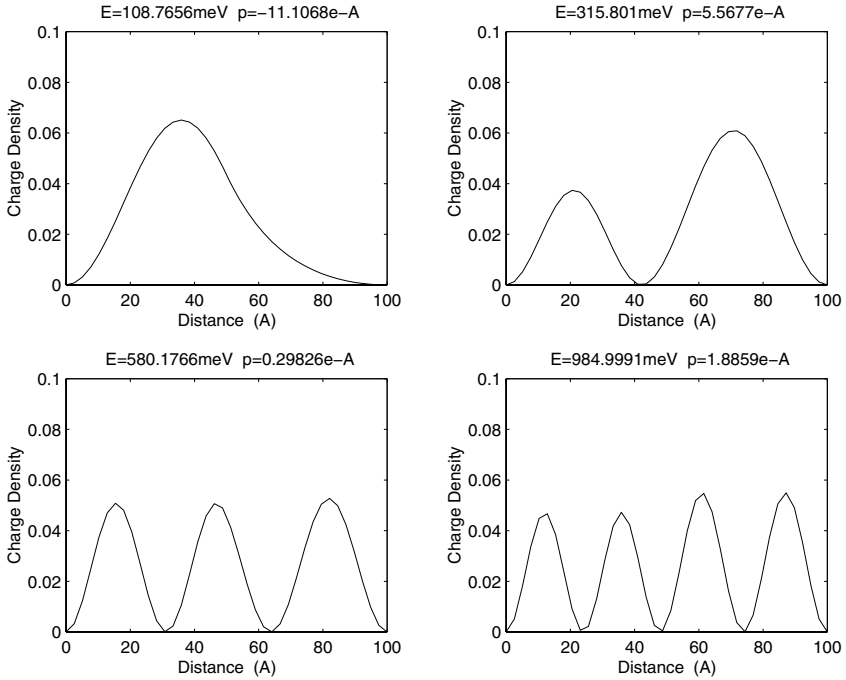


Fig. 2. Charge density distribution for the first four confined states. Results are presented for a quantum dot of width $L_x = 100\text{\AA}$ and a 150 meV high step. The legends on top of each figure give the values of the confined energy and electric dipole.

As one can see, when the dot size is rather small (100\AA), the dipoles due to step-well potentials are also small: $p=11.1\text{ e-\AA}$ for the ground (local) state and $p=5.6\text{ e-\AA}$ for the first excited (global) state. It will be shown further that these values of dipoles do not provide the couplings strong enough for the experimental resolution of different quantum states. For the 200\AA wide quantum dot, the dipoles are 38.8 e-\AA for the ground state, 22.8 e-\AA for the first excited state, and 27.4 e-\AA for the second excited state. Contrary to the previous case, the first excited state here is still a local state which means that its energy ($E=142.3\text{ meV}$) is below the potential step ($E=150\text{ meV}$). The high dipole moments and sufficient intersubband energy separations of the confined levels in the 200\AA wide dot make them attractive for the use as computational basis states. A further increase in the dot size decreases the subband separation too much and is not appropriate for our purposes. It will also be shown in the next section that the confined states of the asymmetric quantum dot are rather long-lived, which is important for computational basis states.

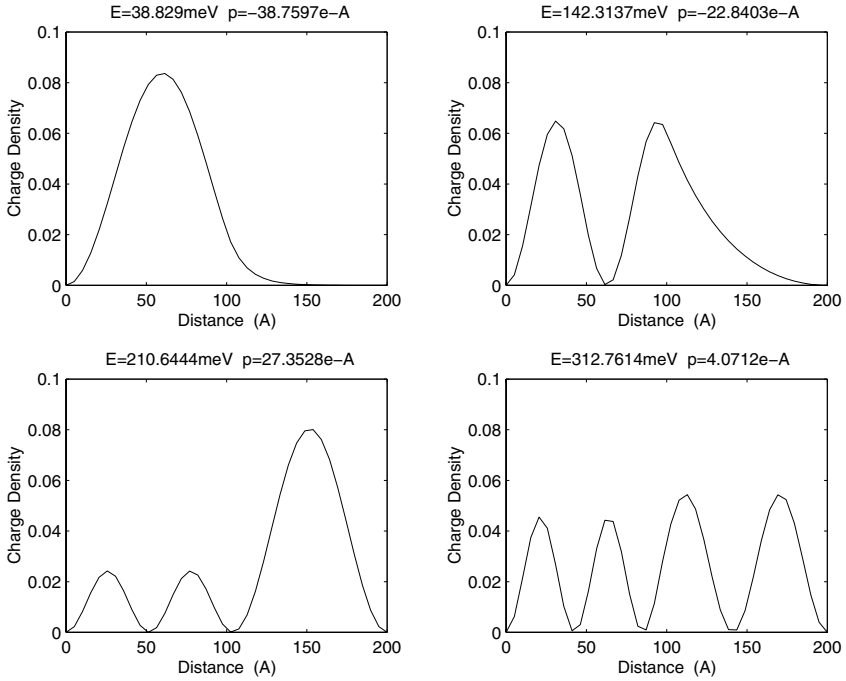


Fig. 3. Charge density distributions for the first four confined states. Results are presented for a quantum dot of width $L_x = 200\text{\AA}$. The legends on top of each figure give the values of the confined energy and electric dipole.

4 Discussion

We have calculated the dipole-dipole interaction coupling constant for the case of two 200\AA wide quantum dots separated by a distance of 100\AA . Fig. 4 shows dipole moments (upper panel) and the coupling constant (lower panel) as functions of the dot size along the z -axis. The height of the potential step is 300 meV . The coupling constant denoted by the solid line is plotted for the gate that utilizes the ground state and the first excited state as the computational basis states. The dashed line corresponds to the gate implemented on the basis of the ground state and second excited state. In the first case (ground state), the coupling constant grows monotonically while in the second case there are well resolved minima which correspond to the dot size and potential profile when the symmetry of the charge density distribution is partially restored. Thus, one has to control precisely the dot size and potential step height when designing the controlled-NOT gate in order to avoid "small-dipole" conditions. Rather large values of the coupling (as compared to the broadening) allows for experimental resolution of different quantum states. Moreover, the condition of Eq. (2) may be met even at room temperature.

A large value of the coupling constant also reduces requirements for the π pulse selectivity. The length of the pulse τ_π has to be in the limits specified by the condition

$$\frac{1}{\omega_d} < \tau_\pi < \tau_{coh}, \quad (4)$$

where τ_{coh} is the quantum coherence time of the system which is limited by spontaneous emission time in our case. For the time allowed for computation (τ_{coh}), we want to send as many π pulses as possible thus increasing the number of computations and prepared quantum states. Since it is difficult to increase upper limit of the inequality (4), we may try to decrease the lower limit. For the asymmetric dot implementation, the inverse of the coupling constant is on the order of $4 \times 10^{-14}\text{ s}$. For comparison, the number obtained in Ref. [4] for the gate based on square potential quantum dots biased with an electric field is on the order of 10^{-12} s . This means that in principle, a femtosecond laser can be used to drive our controlled-NOT gate provided that such laser exists in the working frequency range.

In order to maintain coherence during the computation time and to make τ_{coh} as large as possible, one has to minimize the coupling of the gate to the environment. Particularly for our structure, we have to minimize the probability of spontaneous transitions between the states that serve as the computational basis. Our results indicate that the use of the asymmetric potential instead of biasing electric field may help solve this problem as well. To show this, we calculate the oscillator strength of the transitions using the definition

$$f = \frac{2m\omega}{\hbar} \left| \int_{-L/2}^{L/2} \Psi(z) |\boldsymbol{\nu} \mathbf{r}| \Psi^*(z) dz \right|^2, \quad (5)$$

where ω is the transition frequency, $\boldsymbol{\nu}$ is the unit vector along the direction of the incident photon polarization, \mathbf{r} is the real space radial vector. For simplicity,

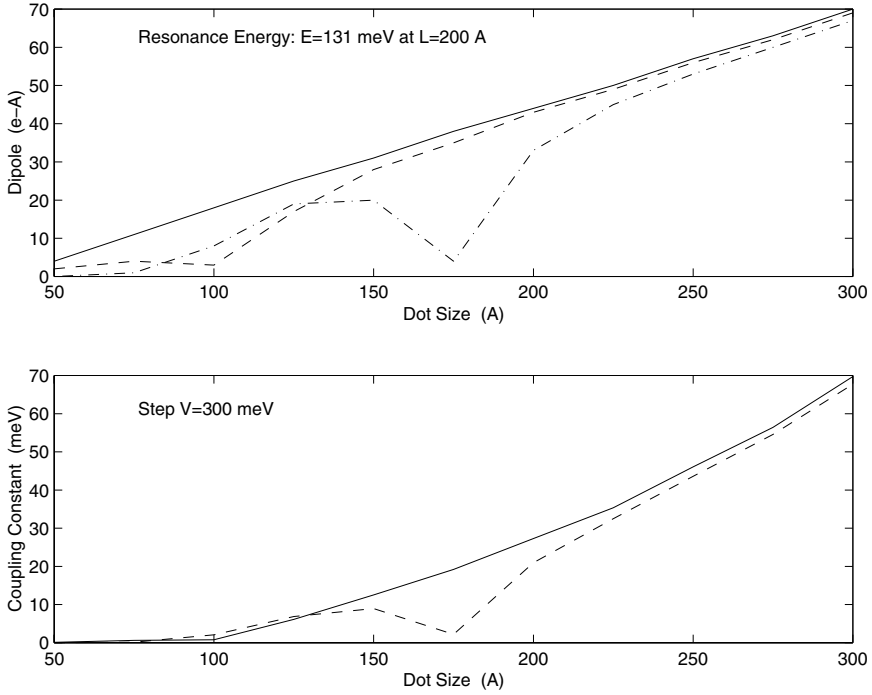


Fig. 4. Upper panel: dipole moments vs. dot size. Solid line, dashed line, and dashed-dotted line correspond to the ground state, the first excited state and the second excited state, respectively. Lower panel: coupling constant of the dipole-dipole interaction vs. dot size. Solid line corresponds to the coupling between the ground state and the first excited state, dashed line corresponds to the coupling between the ground state and the second excited state. The dot separation is fixed at 100 Å

we assume that $\boldsymbol{\nu} = \mathbf{a}_z$ and, thus $\boldsymbol{\nu}\mathbf{r} = z$. The oscillator strength of the transition between the ground and first excited states in a symmetric square potential well (120 Å wide) subjected to an electric field of 36 kV/cm is about 0.95. The same transition in a step potential structure has the value of $f \approx 0.60$. The transition strength between the ground and second excited state in our structure is 0.3 and can be further reduced down to 0.08 if the electric field of 36 kV/cm is applied. From the above arguments, we can conclude that the spontaneous relaxation time $\tau_s (\propto 1/f)$ is bigger in our structure than that in corresponding symmetric dots biased with the electric field. This allows for improvement of the coherence time of the quantum gate by the proper choice of the potential.

5 Conclusions

We propose a solid-state implementation of a quantum controlled-NOT gate based on *coupled asymmetric* quantum dots. The structure can be realized using the state-of-the-art technology. The results of our numerical simulation indicate that owing to the dot asymmetry, the coupling constant of the dipole-dipole interaction can be made very large as compared to the transition line broadening and to the other previously proposed structures [4], [8], [7]. We show that probability of spontaneous emission can be reduced in our gate, thus making it easier to meet the time coherence requirements. Discussion of other issues related to physical implementation of quantum logic gates is also presented. Finally, we argue that our gate can be driven by a femtosecond laser if one is designed for an appropriate frequency range.

Acknowledgement

The work was partially supported by the Jet Propulsion Laboratory. The authors thank Dr. F. Vatan for discussions on quantum computing.

References

- [1] Cirac, J.I., Zoller, P.: Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74** (1995) 4091
- [2] Gershenfeld, N.A., Chuang, I.: Bulk spin-resonance quantum computation. *Science* **275** (1997) 350
- [3] Milburn, G.J.: Quantum optical Fredkin gate. *Phys. Rev. Lett.* **62** (1989) 2124
- [4] Barenco, A., Deutsch, D., Eker, A., Jozsa, R.: Conditional quantum dynamics and logic gates. *Phys. Rev. Lett.* **74** (1995) 4083
- [5] Jones, J.A., Mosca, M.: Implementation of a quantum algorithm to solve Deutsch's problem on a nuclear magnetic resonance quantum computer. LANL preprint quant-ph/980127, 1998
- [6] Unruh, W.G.: Maintaining coherence in quantum computers. *Phys. Rev. A* **51** (1995) 992
- [7] Bandyopadhyay, S., Balandin, A., Roychowdhury, F., Vatan, F.: Nanoelectronic implementation of reversible and quantum logic gates. *Superlattices. Microstruct.* (Special issue in honor of Rolf Landauer) **23** (1998) 445
- [8] Bandyopadhyay, S., Balandin, A.: Quantum dot version of the Toffoli-Fredkin gate and its application in quantum architectures. *Proceed. of the March Meeting of The American Physical Society, Los Angeles, CA, USA, 1998*
- [9] Lloyd, S.: A potentially realizable quantum computer. *Science*. Vol. 261 (1993) 1569
- [10] Wang, K.L., Balandin, A.: (unpublished)
- [11] Yuh, P.F., Wang, K.L.: Large Stark effects for transitions from local states to global states in quantum well structures. *IEEE J. Quantum Electr.* Vol. 25 (1989) 1671
- [12] Yuh, P.F., Wang, K.L.: Optical transitions in step quantum well. *J. Appl. Phys.* **65** (1989) 4377
- [13] Yuh, P.F., Wang, K.L.: Intersubband optical absorption in coupled quantum wells under an applied electric field. *Phys. Rev. B* **38** (1988) 8377

Spatiotemporal Dynamics of Quantum Computing Solid Dipole-Dipole Block Systems

Hideaki Matsueda

Department of Information Science, Kochi University
2-5-1 Akebono-cho, Kochi 780, Japan
matsueda@is.kochi-u.ac.jp

Abstract. This paper enestimates the stability of dipole-dipole interaction in quantum dot array, and proposes a novel solid state quantum CCN (controlled controlled not) gate having a block structure, which is effective to maintain quantum mechanical coherence and reduce both the bit error and the phase error. The spatiotemporal dynamics of quantum computation process involving the quantum entangled pure states is illustrated.

key words: dipole-dipole interaction, quantum dot array, solid block, quantum CCN gate, ensemble cancellation of errors, quantum entangled pure states

1 Introduction

We have been proposing solid state quantum gates, employing a coherent mode generated by the dipole-dipole interaction among ensemble of Frenkel type excitons, each of which being confined in a three dimensional quantum dot [1].

Our treatment of the dipole-dipole interaction is different from previous ones [2][3], especially in that we employed the parallel (analogy to ferromagnetic) dipole-dipole interaction to enhance the coherence, not only to directly execute any particular logic as in the case of ref.[3]. In the next section (§2) of this paper, the stability of this dipole-dipole interaction is estimated energetically.

A pair of the block constitutes a realistic solid state quantum CN (controlled not) gate, having sufficient coherence due to the dipole-dipole interactions among the induced dipole moments, and also due to cancellation of phase fluctuation in the ensemble of quantum dots [1]. Then it is possible to construct universal quantum systems, employing this CN gates and few other kinds of simple gates. In this way, a quantum CCN (controlled controlled not) gate is proposed as an essential part of the universal system for quantum computation.

The spatiotemporal dynamics of the quantum entangled pure states, which is the kernel of the quantum super-parallel computation, is illustrated for the proposed quantum CCN gate of logical block structure. The logical block may work as an artificial molecule having a gigantic dipole moment.

2 Energetics of the Dipole-Dipole Interaction in the Block

2.1 Dipole-Dipole Energy

The array of dipole moments \mathbf{p}_m arranged at positions \mathbf{R}_m impose a resultant Electric field $\mathbf{E}(n)$ on a site at \mathbf{R}_n as given below.

$$\mathbf{E}(n) = \frac{1}{4\pi\epsilon_0} \sum_m \left[\frac{\mathbf{p}_m}{|\mathbf{R}_n - \mathbf{R}_m|^3} - 3 \frac{\mathbf{p}_m \cdot (\mathbf{R}_n - \mathbf{R}_m) (\mathbf{R}_n - \mathbf{R}_m)}{|\mathbf{R}_n - \mathbf{R}_m|^5} \right] \quad (2.1)$$

in SI unit. Then the average energy Q of the total dipole-dipole interaction in a block per single quantum dot may be given as

$$Q = \frac{1}{N} \sum_n \mathbf{p}_n \cdot \mathbf{E}(n) \quad (2.2)$$

where N is the total number of quantum dots in the block.

The dipole-dipole energy of eq. (2.2) is plotted as a function of the separation of quantum dots in Fig.1, for the case of cubic array having identical dipole moments all aligned along z direction. The dipole moment is assumed to be, as an example, that of GaAs; $p = |\mathbf{p}| = ed = 3.5 \times 10^{-10} \text{ e mC}$ [5], with elementary electric charge $e = 1.60219 \times 10^{-19} \text{ C} = 4.80321 \times 10^{-10} \text{ esu}$. The plotting is done for three kinds of blocks, of which dimensions (x, y, z) are (1, 1, 10), (1, 1, 2) and (5, 3, 5).

As is explained in detail elsewhere [1], including this dipole-dipole energy in the starting Hamiltonian, we derived a localized coherent mode below a conventional excitonic band, in a regime where the population difference is below a critical value, which is estimated to be for an example 0.87 (corresponding to 94% excitation) in the case of 2.6nm spacing GaAs quantum dots. In this model, the confinement of the exciton is assumed to be so perfect, that no considerable overlapping of wave function and no tunneling among them is expected. This localized mode may be useful as the excited state of a quantum gate, because of its ferromagnetic type stability.

2.2 Curie Temperature

The field of eq. (2.1) could be thought of as the molecular (or Weiss) field in ferromagnetic materials. For array of identical dipoles aligned along the z direction, the molecular field is given as $|\mathbf{E}(n)| = \lambda N p$, and the Curie-Weiss law $\chi = \frac{C}{T - T_c}$ may be reasonable for the susceptibility χ with a Curie constant $C = \frac{Np^2}{3k}$, $T_c = C\lambda$, and the Boltzmann constant k , following the procedure of

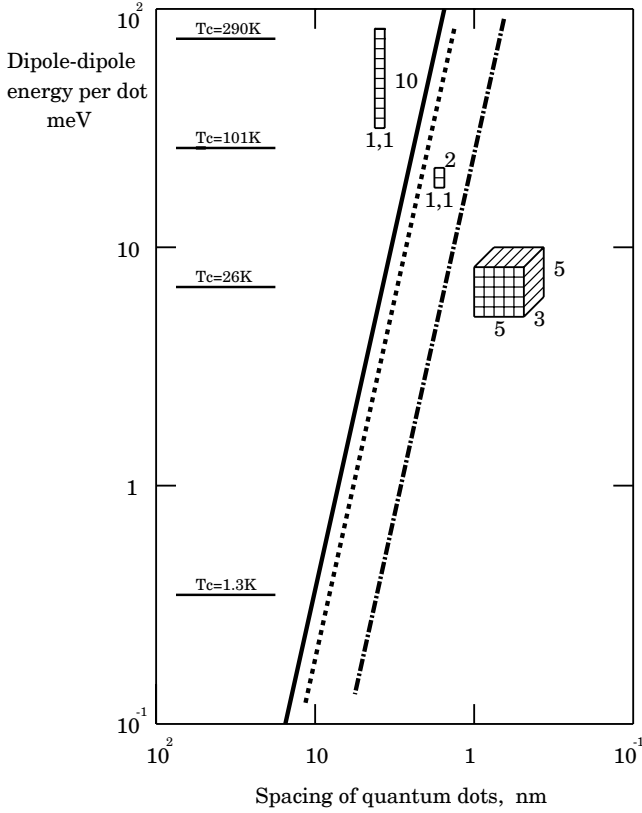


Fig. 1. Dipole-dipole interaction energy per quantum dot as the function of the spacings of quantum dots, for different block dimensions 1-1-10 (solid), 1-1-2 (dotted), and 5-3-5 (dot-dash). The Curie temperature (T_c) for the dipole-dipole energy 75.0, 26.1, 6.8, and 0.35meV are also indicated. The dipole length is assumed to be 3.5×10^{-1} nm as in GaAs.

ferromagnetism [4]. The Weiss constant λ is evaluated using eqs. (2.1) and (2.2) to give the Curie temperature.

$$T_c = -\frac{Q}{3k} \quad (2.3)$$

Typical values of T_c , i.e. 290, 101, 26, and 1.3K are given in Fig.1 for the dipole-dipole energies 75.0, 26.1, 6.8, and 0.35meV respectively.

2.3 Ensemble Stability

It is seen from Fig.1 that the block having the minimum dimension in x-y plane has the largest dipole-dipole energy, ex. (1, 1, 10) block which is an one dimensional array with 10 quantum dots along the z direction. In this case, the parallel

(ferromagnetic type) phase is stable even at room temperature if the quantum dot spacing is less than 1.7nm, and it is stable at 77K and 4K if the spacing is below 2.6nm and 7.1nm respectively. In the (5, 3, 5) block having 5×3 two dimensional array in x-y plane and 5 dots along the z direction, it is stable at 77K and 4K if the spacing is less than 1nm and 2.8nm respectively.

As a matter of course, the stability increases if we employ materials with larger dipole moments; for example in the case of InSb of which dipole moment is approximately 7.14 times larger than that of GaAs [5], the dipole-dipole energy is larger by a factor of 51, resulting in the room temperature T_c at separation of 6.2nm in (1, 1, 10) block, and at 2.5nm in (5, 3, 5) block. This might improve further if some organic materials with gigantic dipole moments are used.

It is also obvious that in this block for the dipole-dipole interaction, the stability does not directly depend on the volume of the block or the number of dipoles, although that is the case in other kinds of memory devices [6]. In our case, the more the side by side dipole arrangement, the less the total dipole-dipole energy gain, because a dipole moment produces a field in the antiparallel direction at the side by side positions. Therefore, the stability depends on the shape of the block rather than the volume of it.

3 Immunity of the Quantum Logical Blocks to Bit and Phase Errors

3.1 Bit Error Immunity

There is a whole bunch of energy levels arising from all the quantum dots in the same block, including energetically degenerate states. Superposition (or intra-block entanglements) of states having same multiple number n of excited sites may be generated in a block having N quantum dots, by an appropriately precision π pulse.

This state

$$\sqrt{\frac{n!(N-n)!}{N!}} \left[|1, 1 \cdots 1, 0, 0, \cdots\rangle + |0, 1, 1 \cdots 1, 0, \cdots\rangle + |0, 0, 1, 1 \cdots 1, 0, \cdots\rangle + \cdots \right] \Rightarrow |\tilde{\mathbf{1}}\rangle \quad (3.1)$$

and the physical ground state

$$|0, 0, 0, \cdots\rangle \Rightarrow |\tilde{\mathbf{0}}\rangle \quad (3.2)$$

should work as a dual basis ($|\tilde{\mathbf{0}}\rangle, |\tilde{\mathbf{1}}\rangle$) for the computation. Thanks to the dipole-dipole interaction, it is obvious from the results of previous section (§2) that these basis states are immune to thermal agitation providing the basic coherence for the quantum computation. First of all, this is powerful to prevent bit errors (amplitude errors) such as $|\tilde{\mathbf{0}}\rangle \rightarrow |\tilde{\mathbf{1}}\rangle$ and $|\tilde{\mathbf{1}}\rangle \rightarrow |\tilde{\mathbf{0}}\rangle$, but this is also effective to avoid phase errors as explained in the next subsection §3.2.

3.2 Phase Error Immunity

It may be possible to generate another orthogonal basis $(|\tilde{\tilde{0}}\rangle, |\tilde{\tilde{1}}\rangle)$ by an Hadamard transformation $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

$$\frac{1}{\sqrt{2}}(|\tilde{0}\rangle + |\tilde{1}\rangle) \Rightarrow |\tilde{\tilde{0}}\rangle \quad (3.3)$$

$$\frac{1}{\sqrt{2}}(|\tilde{0}\rangle - |\tilde{1}\rangle) \Rightarrow |\tilde{\tilde{1}}\rangle. \quad (3.4)$$

This transformation increases the resistance of the coding against phase errors, because a phase error in $(|\tilde{0}\rangle, |\tilde{1}\rangle)$ basis corresponds to a bit error in the $(|\tilde{\tilde{0}}\rangle, |\tilde{\tilde{1}}\rangle)$ basis. as is easily seen from the transformation in eqs. (3.3) and (3.4), for the case of phase error such as $|\tilde{0}\rangle \Rightarrow |\tilde{\tilde{0}}\rangle$ and $|\tilde{1}\rangle \Rightarrow -|\tilde{\tilde{1}}\rangle$, of which consideration may be sufficient for our purpose [7].

There is also cancellation of phase fluctuations in the quantum dot ensemble. For a block with the environment, the phase fluctuations of an off-diagonal element of the density matrix in the $(|\tilde{0}\rangle, |\tilde{1}\rangle)$ basis or even in the $(|\tilde{\tilde{0}}\rangle, |\tilde{\tilde{1}}\rangle)$ basis, should cancel each other, as the number of excited site n increases. This saves the off-diagonal term which generates the entanglement.

The situation is easily seen in the bosonic heat bath model, where the phase fluctuations of the excited state and the ground state both at site k may be represented as ϕ_k and $-\phi_k$ respectively, and their distributions are Gaussian having their means at zero [8]. The total phase fluctuation may be expressed as

$$\begin{aligned} \Phi_i = & -(\phi_1 + \phi_2 + \cdots + \phi_{k-1}) + (\phi_k + \phi_{k+1} + \cdots \\ & + \phi_{k+n-1}) - (\phi_{k+n} + \phi_{k+n+1} + \cdots + \phi_N) \\ \longrightarrow & 0 \text{ for } n, N \rightarrow \infty, \end{aligned} \quad (3.5)$$

for a state i , leading to

$$\begin{aligned} & e^{\Phi_i} |0, 0 \cdots 0, 1, 1 \cdots 1, 0, 0, \cdots\rangle_i \\ &= \left[e^{-i\phi_1} |0_1\rangle \cdots e^{i\phi_k} |1_k\rangle e^{i\phi_{k+1}} |1_{k+1}\rangle \right. \\ & \quad \left. \times \cdots e^{-i\phi_{k+n}} |0_{k+n}\rangle e^{-i\phi_{k+n+1}} |0_{k+n+1}\rangle \cdots \right]_i \\ &= e^{-(\phi_1 + \phi_2 + \cdots + \phi_{k-1}) + (\phi_k + \phi_{k+1} + \cdots + \phi_{k+n-1})} \\ & \quad e^{-(\phi_{k+n} + \phi_{k+n+1} + \cdots + \phi_N)} |0, 0 \cdots 0, 1, 1 \cdots 1, 0, 0, \cdots\rangle_i \\ & \longrightarrow |0, 0 \cdots 0, 1, 1 \cdots 1, 0, 0, \cdots\rangle_i. \end{aligned} \quad (3.6)$$

This may be considered as the ensemble cancellation of random phase fluctuation.

3.3 Error Prevention Methods

There may be two methods to avoid the phase decoherence of encoded states during some significantly complex computation.

One is to divide the total computational algorithm into shorter pieces, in such a way that each piece could be executed by a small block structure. If we denote the statistical variance of the temporal phase fluctuation in each identical quantum dot as σ^2 , then the ensemble average in i th block consisting of N_i quantum dots gives variance $\frac{\sigma^2}{N_i}$. Therefore the whole computing structure of M blocks may suffer from $\sum_{i=1}^M \frac{\sigma^2}{N_i}$, which corresponds to a standard deviation on the order of $\sqrt{\frac{M}{N}}\sigma$.

However, this deviation remains finite, so long as each piece of the algorithm (eventually M) is sufficiently small. The permissible value of $\frac{M}{N}$ depends on the employed error correction scheme and the required accuracy of computation.

The other method is to execute the main part of the computation in the $(|\tilde{0}\rangle, |\tilde{1}\rangle)$ basis given by eqs.(3.3) and (3.4). In this basis, the phase error is suppressed directly by the dipole-dipole interaction among the ensemble of quantum dots in each block, as discussed in §3.1 and §3.2. As a matter of course, some gates or operations should be added to the input and output ports to practice the transformation of eqs. (3.3) and (3.4) between the bases.

4 Quantum Computing Systems

It is possible to construct a quantum computing system out of the blocks given in previous section (§2). As a simple example, a CCN gate is given in Fig.2, which is an essential structure for universal quantum computing system.

Each line of bit a, b, \dots consists of blocks $a_1, a_2, \dots, b_1, b_2, \dots$. A CN operation \hat{P}_{CN} is implemented by the couple of blocks facing each other (i.e. a_3 and b_3 , a_5 and b_5) and a photonic π pulse. The separation of quantum dots is designed to differ line by line, to distinguish individual line from each other photonically [11]. Rotational operations $\hat{R}^{\pi/2}$ and $\hat{R}^{3\pi/2} = \hat{R}^{-\pi/2}$ are also implemented by the couples of blocks (i.e. b_2 and c_2 , b_4 and c_4 , and a_6 and c_6) and a photonic $\frac{\pi}{2}$ pulse and a $\frac{3\pi}{2}$ pulse respectively. The phase difference between lines a, b and line c is adjusted by a pair of controlled phase shifters ($S_{ab}^{\frac{\pi}{2}}$ and $S_{ba}^{\frac{\pi}{2}}$).

The CCN operation of Fig.2 may be described as

$$\hat{P}_{CCN} = \hat{S}_{ba}^{\frac{\pi}{2}} \hat{S}_{ab}^{\frac{\pi}{2}} \hat{R}_{ac}^{\pi/2} \hat{P}_{CNab} \hat{R}_{bc}^{3\pi/2} \hat{P}_{CNab} \hat{R}_{bc}^{\pi/2} \quad (4.1)$$

$$= \hat{S}_{ba}^{\frac{\pi}{2}} \hat{S}_{ab}^{\frac{\pi}{2}} \left[1 + \frac{1}{4}(1 + \sigma_a^z)(1 + \sigma_b^z)(1 - \sigma_c^x) \right], \quad (4.2)$$

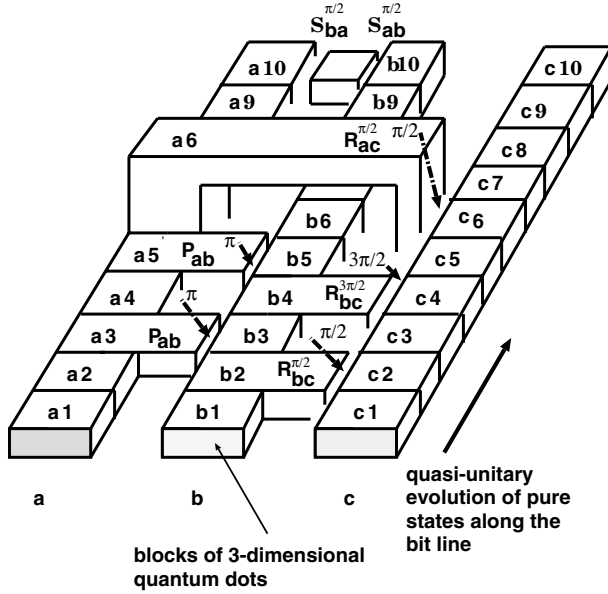


Fig. 2. A solid block CCN gate, combining two CN gates (P_{ab}), two $\frac{\pi}{2}$ controlled rotators ($R_{bc}^{\pi/2}$ and $R_{ac}^{\pi/2}$), one $\frac{3\pi}{2}$ controlled rotator ($R_{bc}^{3\pi/2}$), and a pair of $\frac{\pi}{2}$ controlled phase shifters ($S_{ab}^{\pi/2}$ and $S_{ba}^{\pi/2}$). Arrows with π , $\frac{\pi}{2}$, and $\frac{3\pi}{2}$ denote the respective photonic pulses.

with $\hat{R}^{\pm\pi/2} = \begin{pmatrix} 1 & \mp i \\ \mp i & 1 \end{pmatrix}$, where σ^x , σ^y , σ^z are the Pauli operators, and \pm and \mp correspond to each other.

The non-Boolean superposed state such as $\frac{1}{\sqrt{2}}(|\tilde{0}\rangle + |\tilde{1}\rangle)$ or $\frac{1}{\sqrt{2}}(|\tilde{\tilde{0}}\rangle + |\tilde{\tilde{1}}\rangle)$ may be generated initially by a photonic $\frac{\pi}{2}$ pulse bit by bit, and then conveyed from one block to the other, across different bit lines through the CN gate and the rotational gates, as well as propagated within the same bit line. After an appropriate time period, an expected entangled pure state (spontaneously generated superposition of different product states involving different bit lines) is attained, being specific to the arrangement of the blocks.

The depth and/or width of the quantum dots may be designed to cause a small but sequential decrease of the energy gap, starting from the input port toward the output port, block by block along each bit line. Excitons carrying the dipole moment are transferred in one-way fashion through this biased path, because there exists some dissipative energy loss which deprives the excitons of their energy to climb back the line. Application of microwave pulses with proper energy may work as clock signals. These methods facilitate the directional quasi-

unitary evolution of the whole system, which is needed and sufficient for the quantum computation.

5 Generation of the Quantum Entangled Pure States and Its Spontaneous Accumulation into the Output Port

The spatiotemporal dynamics of the quantum entangled pure states is essential for the quantum super parallel computation. In the quantum CCN gate of Fig.2, it is possible to define a basin for each quantum entanglement generated at a CN gate (P) or at a rotator (R), as indicated in Fig.3 by dotted or dot-dash lines on the top view of the CCN gate, labeling by the times of generation t_1, t_2, \dots , and t_5 . The basin may be thought of as the area that is going to be influenced by a gate starting each entanglement.

All the necessary Boolean algebra are executed by inter-block operations, whereas the purely quantum mechanical superposition and interference of different states will take place in an intra-block manner (within each block). Nonetheless, the quantum interplay of different blocks will be spontaneously realized as the inter-block entanglement via Coulombic interaction, on the way of the unitary (or quasi-unitary) evolution. Then, blocks become incompatible each other, revealing the peculiarity of quantum mechanics. It is no longer possible to measure the states of individual blocks simultaneously, without destroying the states as they are, until the whole thing settle down to yield the computational result in the output port.

The resultant final state is designed to form a pattern of electron excitation with sharply localized peaks at particular blocks in the output port, as the results of the interference of wave functions, which may be achieved by rotational operations embeded in the algorithm such as quantum Fourier transformations. Therefore this result should be read by some kind of photonic probing, involving some auxiliary higher energy states [1].

6 Concluding Remarks

A quantum CCN gate is proposed as the essential elements of the solid state quantum computing system, employing the logical blocks of three dimensional quantum dot array. The spatiotemporal dynamics of the quantum entangled pure states, which is essential for the execution of the quantum super-parallel computation, is illustrated for the proposed quantum CCN gate.

The ferromagnetic type dipole-dipole interaction among the ensemble of induced dipole moments suppresses not only bit errors in the ($|\tilde{\mathbf{0}}\rangle, |\tilde{\mathbf{1}}\rangle$) basis, but also phase errors in ($|\tilde{\mathbf{0}}\rangle, |\tilde{\mathbf{1}}\rangle$) basis. Furthermore, the ensemble cancellation of the phase fluctuation in the block is also expected to improve the phase coherence for the quantum computation.

It is needless to mention that significant technological development is expected, for example precision fabrication of quantum dot array blocks either of

semiconductor or organic compounds, in order to implement the proposed quantum systems for practical computation. Focused photonic pulse technique should also be developed, for an example the scanning near-field method, for the input superposition and the rotational operations.

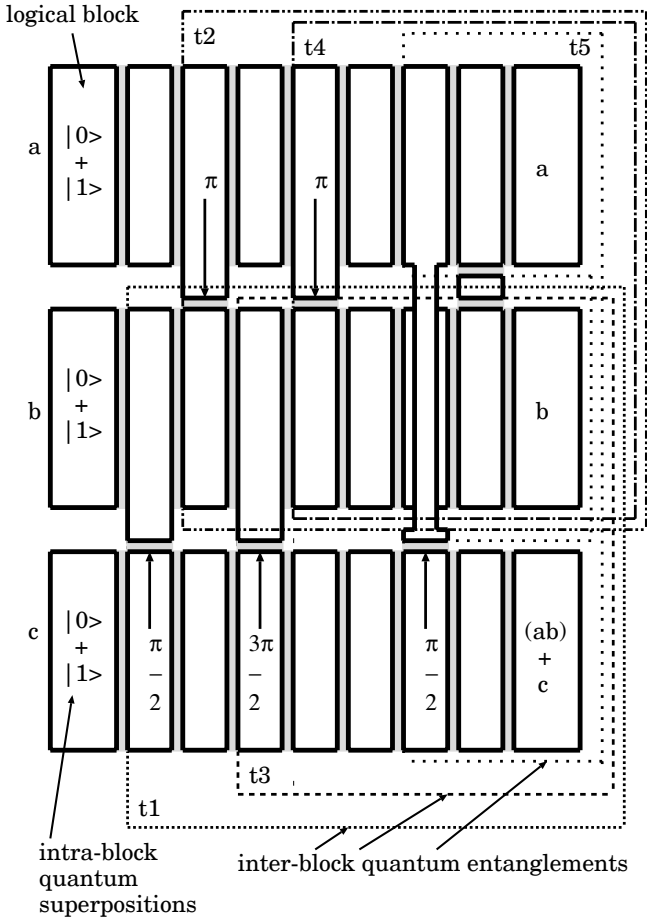


Fig. 3. Spatiotemporal dynamics of quantum entangled pure states in the quantum CCN gate of quantum logical blocks. The basins of 5 entangled pure states generated at 5 different times t_1 , t_2 , \dots are indicated by the dotted or dot-dash lines..

Acknowledgments

This work is partially supported by the Proposal-Based New Industry Creative Type Technology R&D Promotion Program of the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

References

1. H. Matsueda, and S. Takeno, IEICE Trans. Fundamentals Electron., Commun. and Computer Sci., **E79-A**, 1707 (1996); **E80-A**, 1610 (1997); in *Proc. 1st. Int. Conf. on the Theory and Appl. of Cryptology (Pragocrypt'96)*, ed. J. Pribyl (GC UCMP, Praha, 1996) pp.225–233; in *Proc. 4th Workshop on Physics and Computation (PhysComp96)*, eds. T. Toffoli, M. Biafore and J. Leão (New England Complex Systems Institute, Cambridge, Massachusetts, 1996) pp.215–222; H. Matsueda, in *Proc. The European Conference on Circuit Theory and Design (ECCTD'97)*, Budapest (30 Aug. – 3 Sept., 1997) pp.265–270 (*invited paper for the special session "Towards Nanoelectronics"*); in *Proc. the 1st International Conference on Unconventional Models of Computation (UMC'98)*, Auckland, New Zealand, (5–9 Jan., 1998), DMTCS Series, Springer-Verlag.
2. G. Mahler and V. A. Weberruß, *Quantum Networks* (Springer Verlag, Berlin, 1995), Sec. 2.4.2.1; W. G. Teich, K. Obermayer, and G. Mahler, Phys. Rev. B **37**, 8111 (1988); W. G. Teich and G. Mahler, Phys. Rev. A **45**, 3300 (1992).
3. A. Barenco, D. Deutsch, and A. Ekert, Phys. Rev. Lett. **74**, 4083 (1995).
4. C. Kittel, *Introduction to Solid State Physics* 4th ed. (John Wiley, 1971) Chap. 13 and 16.
5. M. Asada and Y. Suematsu, IEEE J. Quantum Electron., **QE-21**, 434 (1985).
6. J. A. Swanson, IBM J., (July, 1960) p.305.
7. A. M. Steane, Phys. Rev. A **54**, 4741 (1996).
8. A. Barenco, A. Ekert, K.-A. Suominen, and P. Törmä, Phys. Rev. A **54**, 139 (1996).

Author Index

- Abrams, D.S. 167
Adami, C. 258, 391
Allahverdyan, A.E. 276, 296
Averin, D.V. 413
- Balandin, A.A. 460
Biham, E. 140
Biham, O. 140
Biron, D. 140
Bowden, C.M. 364
Brassard, G. 1
Briegel, H.-J. 373
Buttler, W.T. 200
Bužek, V. 235
- Castagnoli, G. 189
Catasti, P. 357
Cerf, N.J. 218, 258, 391
Chau, H.F. 314
Chi, D.P. 148
Cirac, J.I. 373
Cleve, R. 61
- van Dam, W. 61
DeVoe, R.G. 438
Despain, A.M. 447
DiVincenzo, D.P. 247
Dowling, J.P. 364
Duan, L.-M. 337
Dür, W. 373
- Ekert, A. 174
van Enk, S.J. 373
- Fijany, A. 10
Franson, J.D. 383
Fuchs, C.A. 247
- Gottesman, D. 302
Grassl, M. 140
Gray, A.G. 113
Grover, L.K. 126
Gulley, M.S. 426
Guo, G.-C. 337
- Hillery, M. 235
Holscheiter, M.H. 426
Hotaling, S.P. 364
Hughes, R.J. 200, 426
- James, D.F.V. 426
Jozsa, R. 103
- Kar, G. 214
Kim, I. 89
Kim, J. 148
Kimble, H.J. 373
Knill, E. 357
Kwiat, P.G. 200, 426
- Laffamme, R. 357
Lamoreaux, S.K. 200, 426
Levitin, L.B. 269
Lidar, D.A. 140
Lloyd, S. 167
Luther, G.G. 200
- Mabuchi, H. 247, 373
Mahler, G. 89
Mariappan, S.V.S. 357
Matsueda, H. 468
Monti, D. 189
Mor, T. 1
Morgan, G.L. 200
Mosca, M. 174
- Nielsen, M. 61
Nordholt, J.E. 200
- Obenland, K.M. 447
Ogburn, R.W. 341
Ozhigov, Y. 152
- Peterson, C.G. 200, 426
Pittman, T.B. 383
Preskill, J. 341
- Roy, S. 214
Roychowdhury, V.P. 325

Saakian, D.B. 276, 296
Sandberg, V.D. 426
Schauer, M.M. 426
Schumacher, B. 285
Simmons, C.M. 200, 426
Smolin, J.A. 247

Tapp, A. 61
Thapliyal, A. 247
Tombesi, P. 402
Tupa, D. 426

Uhlmann, A. 247

Vatan, F. 325
Vitali, D. 402

Wang, K.L. 460
Wang, P.Z. 426
White, A.G. 426
Westmoreland, M. 285
Williams, C.P. 10, 75, 113

Yepez, J. 34

Zak, M. 75, 160
Zoller, P. 373
Zurek, W.H. 357